Universally L^1 good sequences with gaps tending to infinity

Zoltán Buczolich*, Department of Analysis, Eötvös Loránd University, Pázmány Péter Sétány 1/c, 1117 Budapest, Hungary
email: buczo@cs.elte.hu
www.cs.elte.hu/~buczo

February 2, 2008

Abstract

We construct a sequence (n_k) such that $n_{k+1} - n_k \to \infty$ and for any ergodic dynamical system (X, Σ, μ, T) and $f \in L^1(\mu)$ the averages $\lim_{N\to\infty}(1/N)\sum_{k=1}^{N} f(T^{n_k}x)$ converge to $\int_X f d\mu$ for μ almost every x. Since the above sequence is of zero Banach density this disproves a conjecture of J. Rosenblatt and M. Wierdl about the nonexistence of such sequences.

1 Introduction

In [4] it is shown that the sequence $n_k = k^2$, $k = 1, 2, ...$ is L^1-universally bad. This means that for all aperiodic ergodic dynamical systems (X, Σ, μ, T) there exists $f \in L^1(\mu)$ such that the ergodic averages

$$\lim_{N\to\infty} \sum_{k=1}^{N} f(T^{k^2}x) \tag{1}$$

*Research supported by the Hungarian National Foundation for Scientific research T049727.

2000 Mathematics Subject Classification: Primary 37A05; Secondary 28D05, 47A35.

Keywords: ergodic theorem, universally good sequence, Banach density

fail to converge on a set of positive measure. On the other hand, results of Bourgain [1], [2] and [3], imply that (1) converges μ almost everywhere whenever $f \in L^p(\mu)$ with $p > 1$.

When I was working on [4] I learned from M. Keane that it is not known whether there exists a sequence (n_k) such that $n_{k+1} - n_k \to \infty$ and for any $f \in L^1(\mu)$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} f(T^{n_k}x) \tag{2}$$

converges μ almost everywhere. This question is also stated in [7] on p. 64 in the second paragraph of Section 7. A sequence satisfying $n_{k+1} - n_k \to \infty$ is of zero Banach density. In [9] the authors "risk" the following conjecture (see Conjecture 4.1 on p. 74 of [9], here we use slightly different equivalent notation):

Conjecture 1. Suppose that the sequence (n_k) has zero Banach density and let (X, Σ, μ, T) be an aperiodic dynamical system. Then for some $f \in L^1(\mu)$ the averages (2) do not converge almost everywhere.

The purpose of this paper is to show that that there exist universally L^1-good sequences (n_k) for which $n_{k+1} - n_k \to \infty$. A sequence is universally L^1-good if (2) converges μ almost everywhere for any ergodic dynamical system (X, Σ, μ, T) and $f \in L^1(\mu)$. This implies that Conjecture 1 is false. This also provides an explanation why was it so difficult to obtain the result in [4] which states that $n_k = k^2$ is L^1-universally bad.

In this paper, like in [1], we mean by a dynamical system (X, Σ, μ, T) an invertible measure preserving transformation acting on a probability measure space. We also assume that T is aperiodic. By scrutinizing the proof presented in this paper one can see that for our sequence (n_k) the averages (2) converge almost everywhere in ergodic periodic systems as well. The non-invertible case from the point of view of this paper can easily be reduced to the invertible one. Suppose that for a non-invertible aperiodic ergodic dynamical system (X, Σ, μ, T) there exists $f \in L^1(\mu)$ for which (2) diverges when $x \in A \in \Sigma$ and $\mu(A) > 0$. Consider the natural extension $(\widetilde{X}, \widetilde{\Sigma}, \widetilde{\mu}, \widetilde{T})$ of (X, Σ, μ, T) (see [6], Chapter 10, §4., or [8] 1.3.G., p. 13). Then $(\widetilde{X}, \widetilde{\Sigma}, \widetilde{\mu})$ is the inverse limit space obtained from (X, Σ, μ, T). The elements of $\widetilde{X}$ are of the form $\widetilde{x} = (x_0, x_1, ...)$ with $T(x_j) = x_{j-1}$, $j = 1, 2, ...$ The transformation $\widetilde{T}$ is defined so that $\widetilde{T}\widetilde{x} = (Tx_0, Tx_1, ...)$. Then $\widetilde{T}^{-1}\widetilde{x} = (x_1, x_2, ...)$

and by Theorem 1, on p. 241 of [6] $\widetilde{T}$ is an ergodic measure preserving transformation.

Set $\widetilde{A} = \{\widetilde{x} \in \widetilde{X} : x_0 \in A\}$. Then $\widetilde{\mu}(\widetilde{A}) = \mu(A) > 0$. If we set $\widetilde{f}(\widetilde{x}) = f(x_0)$ then $\widetilde{f} \in L^1(\widetilde{\mu})$ and

$$\lim_{N\to\infty}(1/N)\sum_{k=1}^{N}\widetilde{f}(\widetilde{T}^{n_k}\widetilde{x}) = \lim_{N\to\infty}(1/N)\sum_{k=1}^{N}f(T^{n_k}x_0)$$

diverges for all $\widetilde{x} \in \widetilde{A}$. This shows that if (n_k) is L^1-bad for a non-invertible system then it is also bad for a suitable invertible one.

This paper is organized as follows. After this introduction in Section 2 we state Theorem 1 which is the main result of this paper about the existence of universally L^1-good sequences (n_k) with gaps converging to infinity. The proof of Theorem 1 is based on Lemmas 2 and 3. In Lemma 2 we show that the (n_k) averages converge for simple functions, which form a dense subset in L^1. In Lemma 3 a weak $(1,1)$ inequality is established for the maximal operator corresponding to the sequence (n_k).

In Section 3 we define (n_k) by induction. Intervals $[\beta_{m-1}, \beta_m)$ are selected and at the m'th step of our definition we choose the terms of (n_k) in one such interval. One can think of the terms of (n_k) as the union of finitely many arithmetic sequences with those terms deleted which are too close to each other. To be more specific, we choose K_m many different prime numbers $q_{j,m}$ and consider those terms of the set $\{lq_{j,m} : l \in \mathbb{Z}, j = 1, ..., K_m\}$ which are in $[\beta_{m-1}, \beta_m)$ and delete those ones which are too close.

In Section 4 we consider functions on $\mathbb{Z}$ with bounded support. We introduce the operators $\mathcal{B}$ and $\mathcal{B}_0$ with maximal operators $\mathcal{B}^*$ and $\mathcal{B}_0^*$. The maximal inequalities established in this section will be applied in later sections with a fixed $m \in \mathbb{Z}$ for the terms of (n_k) which are in $[\beta_{m-1}, \beta_m)$. The most important result is in Lemma 4 about $\mathcal{B}_0^*$. Lemmas 5 and 6 are mere restatements of well-known maximal inequalities. Lemma 7 contains a not too difficult maximal inequality about the operator $\mathcal{B}^*$.

In Sections 5 and 6 we prove Lemma 3. The second part of the proof of Lemma 3, given in Section 6 is used for the proof of Lemma 2 as well. This means that some estimates and notation introduced here is used only later, in Section 7. This shared proof part explains that instead of using some kind of transference principle why we use directly Kakutani-Rokhlin tower constructions in Sections 5 and 6 to transfer the results from Section 4 to arbitrary dynamical systems. Of course, we also need to "paste" together the

estimates which we obtain for different m's for terms of (n_k) in $[\beta_{m-1}, \beta_m)$. To estimate the (n_k) averages of (2) we represent f as $f = \lambda'(f_{1,m} + f_{2,m} + f_{3,m})$ with $\lambda' \in \mathbb{R}$ and $m \in \mathbb{N}$. In Section 5 we deal with terms involving $f_{2,m}$ and $f_{3,m}$. While the terms involving $f_{1,m}$ are estimated in Section 6. If f is bounded and N is large then in (2) we can replace f by $\lambda' f_{1,m}$ and this is why Section 6 is used in the proof of Lemma 2 as well. In Section 5 during the estimates related to the terms $f_{2,m}$ an operator denoted by B is defined. After the Kakutani-Rokhlin tower construction it turns out that B coincides with $\mathcal{B}$ and the simpler maximal inequality of Lemma 7 can be used to estimate the maximal operators B^* and $\mathcal{B}^*$. It simplifies our work that by (59), $\sum_m f_{2,m} \leq 3f/\lambda'$ and hence $\sum_m f_{2,m} \in L^1$. Unfortunately, it is not always true that $\sum_m f_{1,m} \in L^1$. This is why we need in Section 6 much more sophisticated methods than the ones in Section 5. Here we need to introduce the modified operators B_0 which coincide with $\mathcal{B}_0$ after the Kakutani-Rokhlin tower construction. In this section the more involved Lemma 4 is needed for the estimation of the maximal operators B_0^* and $\mathcal{B}_0^*$.

In Section 7 based on Part 2 of the proof of Lemma 3 we see that for simple functions the (n_k)-averages in (2) do not differ much from the ordinary ergodic averages and hence Birkhoff's Ergodic theorem implies Lemma 2.

2 Main Result

The desired universally L^1-good sequence with gaps tending to infinity will be denoted by (n_k).

We set

$$\overline{N}_a^b = \#\{n_k : n_k \in [a, b)\}.$$

Suppose $f \in L^1(\mu)$. We set

$$A(f, x, N) = \frac{1}{\overline{N}_0^N} \sum_{k=1}^{\overline{N}_0^N} f(T^{n_k} x).$$

We also introduce

$$A^*(f, x) = \sup_{1 \leq N} |A(f, x, N)|.$$

The main result of the paper is the following:

Theorem 1. *There exists a sequence* (n_k) *satisfying* $n_{k+1} - n_k \to \infty$ *(and hence of zero Banach density) which is universally* L^1*-good, that is, for any invertible aperiodic ergodic dynamical system* (X, Σ, μ, T) *and* $f \in L^1(\mu)$ *we have*

$$\lim_{N\to\infty} A(f,x,N) = \lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} f(T^{n_k}x) = \int_X f d\mu, \tag{3}$$

for μ *almost every* $x \in X$.

The proof of Theorem 1 follows from the following two lemmas. The first one yields a dense set in L^1 for which the $A(f,x,N)$ averages converge. A function $f : X \to \mathbb{R}$ is a simple function if it is measurable and its range consists of a finite set.

Lemma 2. *With the assumptions of Theorem 1, for any simple function* f *we have*

$$\lim_{N\to\infty} A(f,x,N) = \int_X f d\mu. \tag{4}$$

The second one gives a weak $(1,1)$ inequality for the maximal operator A^*.

Lemma 3. *With the notation used in Theorem 1 for any* $f \in L^1(\mu)$ *and* $\lambda > 0$ *we have*

$$\mu(\{x : A^*(f,x) > \lambda\}) \le \frac{1000||f||_1}{\lambda}. \tag{5}$$

Proof of Theorem 3. By Lemma 2 there exists a dense set of functions in $L^1(\mu)$ for which $\lim_{N\to\infty} A(f,x,N) = \int_X f d\mu$ holds μ almost everywhere. The weak $(1,1)$ inequality of Lemma 3 then implies the almost everywhere finiteness of the maximal operator $A^*(f,x)$. By Banach's principle the almost everywhere convergence of $A(f,x,N)$ follows for all $f \in L^1(\mu)$ (for the details see [8] 3.2., p. 91). □

For ease of notation, if we write $\int f d\mu$ we always mean $\int_X f d\mu$.

3 Definition of (n_k) and some estimates

We will use some intervals determined by the integers β_m. We set $\beta_{-1} = \beta_0 = 0$ and the positive integers $\beta_1 < ... < \beta_m < ...$ will be defined by induction. In

each block we will use different numbers $q_{j,m}$, $j=1,...,K_m$. These numbers will be different primes if $m>1$. Their product $p_m = q_{1,m}\cdots q_{K_m,m}$ will be called the period used in block m. We suppose that the primes $q_{j,m}$ are approximately the same size, that is,

$$\frac{1}{2} < \frac{q_{j,m}}{q_{j',m}} < 2 \text{ for } j,j' \in \{1,...,K_m\}. \tag{6}$$

We put

$$\widetilde{q}_{j,m} \stackrel{\text{def}}{=} \frac{p_m}{q_{j,m}}, \text{ and } Q(m) \stackrel{\text{def}}{=} \sum_{j=1}^{K_m} \frac{1}{q_{j,m}}.$$

We will use a parameter d_m which will be a lower bound on the distance among the terms of (n_k) belonging to the interval $[\beta_{m-1},\beta_m)$. We suppose that $d_m\to\infty$ and $d_m < q_{j,m}$ for all $j=1,...,K_m$. For example, the choice $d_m=m$ is suitable. The sequence (d_m) will ensure that the gaps between consecutive terms of (n_k) converge to infinity and hence (n_k) will have zero Banach density.

We put $\overline{N}_{-2}=\overline{N}_{-1}=\overline{N}_0=0$ and

$$\overline{N}_m = \overline{N}_0^{\beta_m} = \#\{n_k : n_k \in [0,\beta_m)\}.$$

We will choose our parameters so that $\overline{N}_{m-1}$ is much larger than p_m for $m=2,....$

Next we give the general plan of the definition of our parameters by mathematical induction. There will be several technical assumptions about these parameters introduced later. Here we just want to orientate the reader about what is chosen, when. To start our induction we put $K_1=1, q_{1,1}=1$. Then $p_1=1$ and $Q(1)=1$. At the first step, after $\beta_1>10$ is determined, we will choose the terms of (n_k) in $[\beta_0,\beta_1)$ so that $n_k=k-1$, for $k=1,...,\beta_1$, that is, each integer from $[\beta_0,\beta_1)$ will belong to (n_k).

Suppose for an $m>1$ we have $\beta_{m'-1}$, $K_{m'}$, and $q_{m',j}$ $j=1,...,K_{m'}$ for $m'\le m-1$ and the terms of the sequence n_k which satisfy $n_k<\beta_{m-2}$ are defined. This gives the values of $\overline{N}_{m'}$ for $m'\le m-2$ as well. Choose K_m so that

$$\frac{32}{K_m}\overline{N}_{m-2}10^4\cdot 4^{m+1} < 2^{-(m+1)}. \tag{7}$$

Next, one needs to choose the prime numbers $q_{j,m}>d_m$ so that (6) holds, $p_m=q_{1,m}\cdots q_{K_m,m}>p_{m-1}$, $Q(m)<Q(m-1)$ and

$$\overline{N}_{m-2}\cdot 4\cdot K_m^2\cdot\frac{d_m+1}{\min_{j'}\{q_{j',m}\}} < \frac{1}{200(m+1)}. \tag{8}$$

For $m > 3$ we also set

$$\gamma_m = \frac{1}{2000 \cdot (m+1) \cdot \overline{N}_{m-2}} < \frac{1}{2000 \cdot m \cdot \overline{N}_{m-3}}, \text{ and } \gamma_\beta = \frac{1}{1000}. \tag{9}$$

We put $\gamma_1 = \gamma_2 = \gamma_3 = \frac{1}{8}$.

After the selection of p_m we choose a sufficiently large β_{m-1}.

Later we need for $m = 2, 3, ...$ that by our assumptions

$$1 - \gamma_m > \frac{3}{4} \text{ and } p_{m-1} < p_m < \frac{1}{10^4}(\beta_{m-1} - \beta_{m-2}) < \frac{1}{10^4}\beta_{m-1}. \tag{10}$$

The value of β_{m-1}, and the numbers $q_{j,m-1}$, $j = 1, ..., K_{m-1}$ will determine the terms of (n_k) in $[\beta_{m-2}, \beta_{m-1})$. This will give us the value of $\overline{N}_{m-1}$ as well. We will have several assumptions later about β_{m-1} and $\overline{N}_{m-1}$. One should think of these assumptions that they require that these numbers are much larger than similar parameters with lower indices. For example, we will need that

$$(\sum_{m'=1}^{m-2} \overline{N}_{m'}) \frac{\overline{N}_{m-3}}{\overline{N}_{m-1}} < \frac{1}{3m} \text{ and } p_m < \frac{1}{100}\overline{N}_{m-1} < \frac{1}{100}\beta_{m-1}. \tag{11}$$

In addition, for convenience, we also suppose that

$$p_m \text{ divides } \beta_{m-1}. \tag{12}$$

For ease of notation suppose that β_m and the numbers $q_{j,m}$, $j = 1, ..., K_m$ are given for an $m > 2$. Next we discuss how these numbers determine (n_k) in $[\beta_{m-1}, \beta_m)$ for $m > 1$. According to (12), p_m and hence all $q_{j,m}$ divide β_{m-1}. Set

$$\Lambda_{j,m,0} = \{lq_{j,m} : l \in \mathbb{Z}\} \cap [\beta_{m-1}, \beta_m).$$

If we take a union of the sets $\Lambda_{j,m,0}$ for $j = 1, ..., K_m$ then some elements might be closer than d_m. So we need to remove these points. First set

$$\Lambda^-_{j,m,0} = \left\{ n \in \Lambda_{j,m,0} : \exists\, n' \in \bigcup_{j'=1,\ j' \neq j}^{K_m} \Lambda_{j',m,0},\ |n' - n| \leq d_m \right\}$$

then put

$$\Lambda_m = \bigcup_{j=1}^{K_m} \Lambda_{j,m,0} \setminus \Lambda^-_{j,m,0}.$$

Since β_{m-1} belongs to all $\Lambda_{j,m,0}$ we have

$$[\beta_{m-1}, \beta_{m-1} + d_m) \cap \Lambda_m = \emptyset. \tag{13}$$

We define the terms of (n_k) so that $n_{k-1} < n_k$ and $\{n_k\}_{k=1}^{\infty} \cap [\beta_{m-1}, \beta_m) = \Lambda_m \cap [\beta_{m-1}, \beta_m)$. Therefore, letting $\Lambda_1 = [\beta_0, \beta_1)$ we have $\{n_k : k = 1, ...\} = \cup_{m\in\mathbb{N}}\Lambda_m$ and the spacing of at least d_m among the elements of each Λ_m plus (13) ensures that $n_{k+1} - n_k \to \infty$ as $k \to \infty$.

Suppose $[n', n' + p_m) \subset [\beta_{m-1}, \beta_m)$. Then

$$\#(\Lambda_{j,m,0} \cap [n', n' + p_m)) = \frac{p_m}{q_{j,m}} = \widetilde{q}_{j,m},$$

and

$$\#(\Lambda_m \cap [n', n' + p_m)) \le \sum_{j=1}^{K_m} \#(\Lambda_{j,m,0} \cap [n', n' + p_m)) = p_m Q(m). \tag{14}$$

For $j' \neq j$ set

$$\Lambda^{-}_{j,m,0,j'} = \{lq_{j,m} \in \Lambda_{j,m,0} : \exists l' \in \mathbb{Z},\ \text{such that } |lq_{j,m} - l'q_{j',m}| \le d_m\}.$$

If $j' \neq j$ then $q_{j,m}$ and $q_{j',m}$ are relatively prime. Modulo $q_{j',m}$ the numbers $lq_{j,m}$, $l = 0, ..., q_{j',m} - 1$ hit each residue class exactly once. Hence, out of these $2d_m + 1$ are not farther than d_m from 0 modulo $q_{j',m}$. Thus, for each $j' \neq j$ out of the $\widetilde{q}_{j,m}$ many elements of $\Lambda_{j,m,0} \cap [n', n' + p_m)$ we need to delete less than $2(d_m+1)\widetilde{q}_{j,m}/q_{j',m}$ many for being too close to an element of $\Lambda_{j',m,0}$. We have a lower estimate

$$\begin{aligned}\#(\Lambda_m \cap [n', n' + p_m)) &> \sum_{j=1}^{K_m} \widetilde{q}_{j,m}\left(1 - \sum_{j'\neq j} 2(d_m + 1) \cdot \frac{1}{q_{j',m}}\right) > \\ &\left(\sum_{j=1}^{K_m} \widetilde{q}_{j,m}\right)\left(1 - K_m \frac{2(d_m + 1)}{\min_{j'}\{q_{j',m}\}}\right) = p_m Q(m)\left(1 - K_m \frac{2(d_m + 1)}{\min_{j'}\{q_{j',m}\}}\right) \\ &> (1 - \gamma_m) p_m Q(m),\end{aligned} \tag{15}$$

where, taking into consideration (9), the last inequality for $m > 3$ needs the assumption

$$\frac{1}{2000(m + 1)\overline{N}_{m-2}} > K_m \frac{2(d_m + 1)}{\min_{j'}\{q_{j',m}\}} \tag{16}$$

about our initial parameters which can be achieved by choosing the $q_{j,m}$'s sufficiently large. For $m = 2, 3$ one needs to put $1/8$ to the left-hand side of (16) when this assumption is made. Combining (14) and (15) one can see that in any "period" $[n', n' + p_m) \subset [\beta_{m-1}, \beta_m)$ the sequence (n_k) has a little less than $p_m Q(m)$ many terms, and $Q(m)$ approximately equals the density of this sequence here. This can be reformulated as

$$1 > \frac{\#(\Lambda_m \cap [n', n' + p_m))}{p_m Q(m)} > 1 - \gamma_m, \tag{17}$$

or as

$$1 > \frac{\#(\Lambda_m \cap [n', n' + p_m))}{\sum_{j=1}^{K_m} \#(\Lambda_{j,m,0} \cap [n', n' + p_m))} > 1 - \gamma_m. \tag{18}$$

Later we need some assumptions and estimations about our parameters. In the rest of this section we give some of these, not too difficult, but rather technical calculations.

We can choose our initial parameters so that for all $m > 0$ with γ_β defined in (9) we have

$$\beta_{m-1} + 2p_m < \frac{\gamma_\beta}{2}\beta_m. \tag{19}$$

This implies

$$\beta_m(1 - \gamma_\beta) < (\beta_m - \beta_{m-1} - 2p_m). \tag{20}$$

Set $P_m = \lfloor \frac{\beta_m - \beta_{m-1}}{p_m} \rfloor$. By (17)

$$1 > \frac{\#(\Lambda_m \cap [\beta_{m-1}, \beta_{m-1} + P_m p_m))}{P_m p_m Q(m)} > 1 - \gamma_m, \tag{21}$$

and by (14) we also have

$$\#(\Lambda_m \cap [\beta_{m-1}, \beta_{m-1} + P_m p_m)) \le \overline{N}_{\beta_{m-1}}^{\beta_m} < \tag{22}$$

$$\#(\Lambda_m \cap [\beta_{m-1}, \beta_{m-1} + P_m p_m)) + p_m Q(m).$$

We need more estimates of $\overline{N}_{\beta_{m-1}}^{\beta_m}$ from above, and from below. By (21) and (22)

$$\overline{N}_{\beta_{m-1}}^{\beta_m} > (1 - \gamma_m) P_m p_m Q(m) = \tag{23}$$

$$(1 - \gamma_m) \left\lfloor \frac{\beta_m - \beta_{m-1}}{p_m} \right\rfloor p_m Q(m) >$$

(using (20))

$$(1-\gamma_m)((\beta_m-\beta_{m-1})-p_m)Q(m) > (1-\gamma_m)(1-\gamma_\beta)\beta_m Q(m),$$

on the other hand,

$$\overline{N}_{\beta_{m-1}}^{\beta_m} < P_m p_m Q(m) + p_m Q(m) = (P_m+1)p_m Q(m) < \tag{24}$$

(using (10))

$$(\beta_m-\beta_{m-1}+p_m)Q(m) < \beta_m Q(m).$$

We suppose that an m_0 is given and $\beta_{m_0-1} < N \le \beta_{m_0}$. Set $P_{m_0,N} = \lfloor \frac{N-\beta_{m_0-1}}{p_{m_0}} \rfloor$. By (17)

$$1 \ge \frac{\#(\Lambda_{m_0}\cap[\beta_{m_0-1},\beta_{m_0-1}+P_{m_0,N}p_{m_0}))}{P_{m_0,N}p_{m_0}Q(m_0)} > 1-\gamma_{m_0}, \tag{25}$$

where we regard $0/0=1$ by definition. We also have

$$\begin{aligned}\#(\Lambda_{m_0}\cap[\beta_{m_0-1},\beta_{m_0-1}+P_{m_0,N}p_{m_0})) \le \overline{N}_{\beta_{m_0}-1}^{N} < \\ \#(\Lambda_{m_0}\cap[\beta_{m_0-1},\beta_{m_0-1}+P_{m_0,N}p_{m_0})) + p_{m_0}Q(m_0),\end{aligned} \tag{26}$$

which implies

$$\begin{aligned}\overline{N}_{\beta_{m_0-1}}^{N} \ge (1-\gamma_{m_0})P_{m_0,N}p_{m_0}Q(m_0) > \\ (1-\gamma_{m_0})(N-\beta_{m_0-1}-p_{m_0})Q(m_0),\end{aligned} \tag{27}$$

and, on the other hand

$$\overline{N}_{\beta_{m_0-1}}^{N} < (P_{m_0,N}+1)p_{m_0}Q(m_0) \le (N-\beta_{m_0-1}+p_{m_0})Q(m_0). \tag{28}$$

To estimate $\overline{N}_0^N$ from below we combine (23) for $m<m_0$ with (27)

$$\overline{N}_0^N = \sum_{m=1}^{m_0-1} \overline{N}_{\beta_{m-1}}^{\beta_m} + \overline{N}_{\beta_{m_0-1}}^{N} \ge \tag{29}$$

$$\sum_{m=1}^{m_0-1}(1-\gamma_m)(\beta_m-\beta_{m-1}-p_m)Q(m) + (1-\gamma_{m_0})(N-\beta_{m_0-1}-p_{m_0})Q(m_0) >$$

(using (10), (27) and $Q(m-1)\ge Q(m)$, $m=2,3,...$)

$$\frac{3}{4}\Bigg(\sum_{m=1}^{m_0-1}\frac{99}{100}(\beta_m-\beta_{m-1})Q(m)+(N-\beta_{m_0-1})Q(m_0)-\frac{1}{100}(\beta_{m_0-1}-\beta_{m_0-2})Q(m_0)\Bigg)$$

$$> \frac{3}{4}Q(m_0)\frac{98}{100}N > \frac{3}{5}Q(m_0)N.$$

4 Functions on $\mathbb{Z}$

Assume $\phi : \mathbb{Z} \to \mathbb{C}$ is of finite support and $|\phi| \leq M$. For ease of notation in this section we drop the subscript m corresponding to the m'th step of the definition of (n_k). So we assume that $q_1, ..., q_K$ are different primes and $p = q_1 \cdots q_K$. We will consider the $[(t-1)p, tp) \cap \mathbb{Z}$ "grid intervals". Given $n \in \mathbb{Z}$ choose $t(n)$ such that $n \in [(t(n)-1)p, t(n)p)$. (In case there is a possibility of misunderstanding we will write $t \cdot (n+2)$ for the product of t and $(n+2)$ and $t(n+2)$ for the function $t(.)$ evaluated at $n+2$.) For any $t \in \mathbb{Z}$ set $\phi_{t,0}(n) = \phi(n - (t(n)-t)p)$. This function is periodic by p and coincides with ϕ on $[(t-1)p, tp) \cap \mathbb{Z}$, hence

$$\phi_{t(n),0}(n) = \phi(n). \tag{30}$$

We also put

$$\overline{\phi}_0(n) = \frac{1}{p} \sum_{k=(t(n)-1)p}^{t(n)p-1} \phi(k) = \frac{1}{p} \sum_{k=0}^{p-1} \phi_{t(n),0}(k),$$

so $\overline{\phi}_0(n)$ is the average of ϕ on the interval $[(t(n)-1)p, t(n)p)$. Observe that

$$\sum_{n=-\infty}^{\infty} \overline{\phi}_0(n) = \sum_{n=-\infty}^{\infty} \phi(n). \tag{31}$$

Set

$$t_0(n,N) = \left\lfloor \frac{n}{p} \right\rfloor + 1 = t(n),\ t_1(n,N) = \left\lfloor \frac{n+N}{p} \right\rfloor + 1, \text{ and} \tag{32}$$

$$N' = t_1(n,N) - t_0(n,N) + 1.$$

For given n and N set

$$I(n,N) = \Big[(t_0(n,N)-1)p - n, t_1(n,N)p - n\Big) \cap \mathbb{Z}, \tag{33}$$

$$\nu(n,N) = \#I(n,N) = N' \cdot p,\ \nu(n,N,j) = \nu(n,N)/q_j = N'\widetilde{q}_j. \tag{34}$$

Clearly, $\nu(n,N) \leq N + p$. We keep assumption (6), that is,

$$1/2 < q_j/q_{j'} < 2, \text{ for any } j, j'. \tag{35}$$

We introduce the operators

$$\mathcal{B}(\phi, n, N, j) = \frac{1}{\nu(n, N, j)} \sum_{lq_j \in I(n,N)} \phi(n + lq_j), \tag{36}$$

$$\mathcal{B}(\phi, n, N) = \frac{\sum_{j=1}^{K} \nu(n, N, j)\mathcal{B}(\phi, n, N, j)}{\sum_{j=1}^{K} \nu(n, N, j)},$$

and their "modified versions"

$$\mathcal{B}_0(\phi, n, N, j) = \frac{1}{\nu(n, N, j)} \sum_{t=t_0(n,N)}^{t_1(n,N)} \left| \sum_{lq_j + n \in [(t-1)p, tp)} \phi(n + lq_j) - \overline{\phi}_0(n + lq_j) \right|, \tag{37}$$

$$\mathcal{B}_0(\phi, n, N) = \frac{\sum_{j=1}^{K} \nu(n, N, j)\mathcal{B}_0(\phi, n, N, j)}{\sum_{j=1}^{K} \nu(n, N, j)}.$$

Using (33-35) it is not difficult to see that

$$|\mathcal{B}_0(\phi, n, N)| \leq \frac{2}{K} \sum_{j=1}^{K} |\mathcal{B}_0(\phi, n, N, j)|. \tag{38}$$

The corresponding maximal operators are defined as

$$\mathcal{B}_0^*(\phi, n, j) = \sup_{N \geq 1} |\mathcal{B}_0(\phi, n, N, j)|, \text{ and } \mathcal{B}_0^*(\phi, n) = \sup_{N \geq 1} |\mathcal{B}_0(\phi, n, N)|.$$

One of the main tools we will use later is the next lemma.

Lemma 4. *For any $\phi : \mathbb{Z} \to \mathbb{C}$ of finite support, which is bounded by M we have*

$$||\mathcal{B}_0^*(\phi, .)||_{\ell^2} \leq \frac{32}{K} M ||\phi||_{\ell^1}. \tag{39}$$

The most useful ingredient in (39) will be K in the denominator of the right-hand side.

In some estimates Lemma 4 will be used instead of the usual maximal inequality (Lemma 3.5, p. 62 of [9]):

Lemma 5. *For all $\phi : \mathbb{Z} \to \mathbb{C}$ of finite support for all $\lambda > 0$,*

$$\#\left\{ n : \sup_{N \geq 1} \left| \frac{1}{N} \sum_{k=0}^{N-1} \phi(n + k) \right| > \lambda \right\} \leq \frac{2}{\lambda} ||\phi||_{\ell^1}. \tag{40}$$

We will also need the strong maximal inequality from Lemma 4.4 of [9]. Here we give only the special case of this lemma concerning ℓ^2 norms, and use slightly different notation.

Lemma 6. *For any $\phi : \mathbb{Z} \to \mathbb{C}$ of finite support*

$$\left|\left| \sup_{N>0} \frac{1}{N} \sum_{k=1}^{N} \phi(n+k) \right|\right|_{\ell^2} \leq 2||\phi||_{\ell^2}. \tag{41}$$

We recall a few basic facts about discrete Fourier transforms.

For ease of notation we put $e(x) = \exp(2\pi i x)$.

Given a function $\phi : \{0, ..., p-1\} \to \mathbb{C}$ we set

$$\widehat{\phi}(\frac{b}{p}) = \frac{1}{p} \sum_{n=0}^{p-1} \phi(n) e(-\frac{nb}{p}) \text{ for } b = 0, ..., p-1. \tag{42}$$

Since $e(x)$ is periodic by one the definition of $\widehat{\phi}(b/p)$ can be extended for all $b \in \mathbb{Z}$.

The inverse Fourier transform of $\psi : \{0, \frac{1}{p}, ..., \frac{p-1}{p}\} \to \mathbb{C}$ is

$$\check{\psi}(n) = \sum_{b=0}^{p-1} \psi(\frac{b}{p}) e(n\frac{b}{p}) \text{ for } n = 0, ..., p-1. \tag{43}$$

The way $\widehat{\phi}$ and $\check{\psi}$ are normalized differ in some treatments, sometimes the factor $1/p$ is used in the definition of $\check{\psi}$ and sometimes factors of $1/\sqrt{p}$ are used in both definitions of $\widehat{\phi}$ and $\check{\psi}$. With our choice of normalization Parseval's theorem can be stated as

$$\frac{1}{p} \sum_{n=0}^{p-1} |\phi(n)|^2 = \sum_{b=0}^{p-1} |\widehat{\phi}(\frac{b}{p})|^2. \tag{44}$$

Next we turn to the proof of Lemma 4.

Proof of Lemma 4. Set

$$\phi_{t,0,j}(n) = \frac{1}{\widetilde{q}_j} \sum_{k=0}^{\widetilde{q}_j - 1} \phi_{t,0}(n + kq_j). \tag{45}$$

This function is periodic by q_j while $\phi_{t,0}$ is periodic by $p = q_j \widetilde{q}_j$. The Fourier transform of $\phi_{t,0}$ is

$$\widehat{\phi}_{t,0}(\frac{b}{p}) = \frac{1}{p} \sum_{n=0}^{p-1} \phi_{t,0}(n) e(-\frac{nb}{p}),$$

while the Fourier transform of $\phi_{t,0,j}$ equals

$$\widehat{\phi}_{t,0,j}(\frac{b}{p}) = \widehat{\phi}_{t,0}(\frac{b}{p}) \frac{1}{\widetilde{q}_j} \sum_{k=0}^{\widetilde{q}_j - 1} e(\frac{kq_j b}{p}) = \begin{cases} \widehat{\phi}_{t,0}(\frac{b}{p}), & \text{if} \quad \widetilde{q}_j | b; \\ 0, & \text{if} \quad \widetilde{q}_j \not| b. \end{cases} \tag{46}$$

Recall that q_j and $q_{j'}$ are different primes when $j \neq j'$. Hence $0 < b = r\widetilde{q}_j = rp/q_j = r'\widetilde{q}_{j'} = r'p/q_{j'} < p$ with integers $0 < r < q_j$ and $0 < r' < q_{j'}$ would imply $rq_{j'} = r'q_j$, but this is impossible. Since $\widehat{\phi}_{t,0,j}$ is periodic by one from this it follows that for $b/p \neq 0$ (modulo one) and $j \neq j'$ we have

$$\widehat{\phi}_{t,0,j}(\frac{b}{p}) \widehat{\phi}_{t,0,j'}(\frac{b}{p}) = 0. \tag{47}$$

Suppose $n \in [(t-1)p, tp)$. Then $\widehat{\phi}_{t,0,j}(0) = \widehat{\phi}_{t,0}(0) = \overline{\phi}_0(n)$. Set

$$\phi_{t,0,j,-}(n) = \phi_{t,0,j}(n) - \widehat{\phi}_{t,0,j}(0) = \phi_{t,0,j}(n) - \overline{\phi}_0(n),$$

and

$$\phi_{t,0,-}(n) = \phi_{t,0}(n) - \widehat{\phi}_{t,0}(0) = \phi_{t,0}(n) - \overline{\phi}_0(n) = \phi(n) - \overline{\phi}_0(n)$$

where, again, in the first display the last equality and in the second display the last two equalities hold when $n \in [(t-1)p, tp)$, that is, $t = t(n)$ while the other equalities make sense for other n's as well.

It is also clear that

$$\widehat{\phi}_{t,0,-}(\frac{b}{p}) = \widehat{\phi}_{t,0}(\frac{b}{p}) \text{ if } b/p \neq 0 \bmod 1, \text{ and } \widehat{\phi}_{t,0,-}(0) = 0. \tag{48}$$

We also put

$$\phi_{0,j,-}(n) = \phi_{t(n),0,j,-}(n), \text{ and } \phi^*_{0,j,-}(n) = \sup_{N>0} \frac{1}{N} \sum_{k=0}^{N-1} |\phi_{0,j,-}(n+kp)|.$$

By the strong maximal inequality (Lemma 6) used on $n + p\mathbb{Z}$ instead of $\mathbb{Z}$,

$$\sum_{k=-\infty}^{\infty} |\phi^*_{0,j,-}(n + kp)|^2 \leq 2 \sum_{k=-\infty}^{\infty} |\phi_{0,j,-}(n + kp)|^2.$$

Therefore,

$$\sum_{n=0}^{p-1} \sum_{k=-\infty}^{\infty} |\phi^*_{0,j,-}(n + kp)|^2 = \sum_{n=-\infty}^{\infty} |\phi^*_{0,j,-}(n)|^2 \leq 2 \sum_{n=-\infty}^{\infty} |\phi_{0,j,-}(n)|^2. \tag{49}$$

By Parseval's theorem and (48)

$$\frac{1}{p} \sum_{n=0}^{p-1} |\phi_{0,j,-}(n + (t-1)p)|^2 = \sum_{b=1}^{p} |\widehat{\phi}_{t,0,j}(\frac{b}{p})|^2,$$

where we recall that $\widehat{\phi}_{t,0,j,-}(0) = 0$, so this term is left out from the summation on the right-hand side of the above formula. It was the main motivation for introducing the operators $\mathcal{B}_0$, functions $\phi_{t,0,j,-}$ and $\phi_{t,0,-}$. Thus, keeping t fixed

$$\frac{1}{K} \sum_{j=1}^{K} \sum_{n=0}^{p-1} |\phi_{0,j,-}(n + (t-1)p)|^2 = \frac{p}{K} \sum_{j=1}^{K} \sum_{b=1}^{p} |\widehat{\phi}_{t,0,j}(\frac{b}{p})|^2 \leq \tag{50}$$

(using (30), (45-48), and Parseval's theorem)

$$\frac{p}{K} \sum_{b=1}^{p} |\widehat{\phi}_{t,0,-}(\frac{b}{p})|^2 = \frac{1}{K} \sum_{n=0}^{p-1} |\phi_{t,0,-}(n + (t-1)p)|^2 =$$
$$\frac{1}{K} \sum_{n=0}^{p-1} |\phi(n+(t-1)p) - \overline{\phi}_0(n + (t-1)p)|^2.$$

Next we show that

$$\phi^*_{0,j,-}(n) = \mathcal{B}^*_0(\phi, n, j). \tag{51}$$

By (32-34), $\nu(n,N)=N'p$, $\nu(n,N,j)=N'\widetilde{q}_j$, and by its definition

$$\mathcal{B}_0(\phi,n,N,j)=\frac{1}{\nu(n,N,j)}\sum_{t=t_0(n,N)}^{t_1(n,N)}\left|\sum_{lq_j+n\in[(t-1)p,tp)}\phi(n+lq_j)-\overline{\phi}_0(n+lq_j)\right|= \tag{52}$$

$$\frac{1}{N'}\sum_{k=0}^{N'-1}\left|\frac{1}{\widetilde{q}_j}\sum_{lq_j+n\in[(t(n+kp)-1)p,t(n+kp)p)}(\phi(n+lq_j)-\overline{\phi}_0(n+kp))\right|=$$

$$\frac{1}{N'}\sum_{k=0}^{N'-1}|\phi_{t(n+kp),0,j}(n+kp)-\overline{\phi}_0(n+kp)|=\frac{1}{N'}\sum_{k=0}^{N'-1}|\phi_{0,j,-}(n+kp)|. \tag{53}$$

Taking supremum with respect to N in (52), which means taking supremum with respect to N' in (53), we obtain (51).

Clearly, by (33-35) and (37)

$$|\mathcal{B}_0^*(\phi,n)|\leq\frac{\sum_{j=1}^K\nu(n,N,j)|\mathcal{B}_0^*(\phi,n,j)|}{\sum_{j=1}^K\nu(n,N,j)}\leq\frac{2}{K}\sum_{j=1}^K|\mathcal{B}_0^*(\phi,n,j)|.$$

Using this, (51) and the Cauchy-Schwarz inequality

$$|\mathcal{B}_0^*(\phi,n)|^2\leq\frac{4}{K^2}\left(\sum_{j=1}^K|\mathcal{B}_0^*(\phi,n,j)|\right)^2\leq$$

$$\frac{4}{K^2}\left(\sqrt{K}\sqrt{\sum_{j=1}^K|\mathcal{B}_0^*(\phi,n,j)|^2}\right)^2\leq$$

$$\frac{4}{K}\sum_{j=1}^K|\mathcal{B}_0^*(\phi,n,j)|^2=\frac{4}{K}\sum_{j=1}^K|\phi_{0,j,-}^*(n)|^2.$$

Therefore, using (49)

$$\sum_{n=-\infty}^{\infty}|\mathcal{B}_0^*(\phi,n)|^2\leq\frac{4}{K}\sum_{j=1}^K\sum_{n=-\infty}^{\infty}|\phi_{0,j,-}^*(n)|^2\leq\frac{8}{K}\sum_{j=1}^K\sum_{n=-\infty}^{\infty}|\phi_{0,j,-}(n)|^2=$$

$$\frac{8}{K}\sum_{j=1}^{K}\sum_{t=-\infty}^{\infty}\sum_{n=0}^{p-1}|\phi_{0,j,-}(n+(t-1)p)|^2 \leq$$

(using (50))

$$\sum_{t=-\infty}^{\infty}\frac{8}{K}\sum_{n=0}^{p-1}|\phi_{t,0,-}(n+(t-1)p)|^2 = \frac{8}{K}\sum_{n=-\infty}^{\infty}|\phi(n)-\overline{\phi}_0(n)|^2 \leq$$

(using (31))

$$\frac{16}{K}\sum_{n=-\infty}^{\infty}(|\phi(n)|^2+|\overline{\phi}_0(n)|^2) \leq \frac{16}{K}M\sum_{n=-\infty}^{\infty}(|\phi(n)|+|\overline{\phi}_0(n)|) = \frac{32}{K}M||\phi||_{\ell^1}.$$

$\square$

We also need a weak $(1,1)$ inequality for the operator $\mathcal{B}^*$ which is defined as follows:

$\mathcal{B}^*(\phi,n) = \sup_{N>0}|\mathcal{B}(\phi,n,N)|$ and we also use $\mathcal{B}^*(\phi,n,j) = \sup_{N>0}|\mathcal{B}(\phi,n,N,j)|$.

By using the definition of $\mathcal{B}(\phi,n,N)$ and (35) it is easy to see that if $\phi \geq 0$ then

$$\mathcal{B}(\phi,n,N) \leq \frac{2}{K}\sum_{j=1}^{K}\mathcal{B}(\phi,n,N,j). \tag{54}$$

Lemma 7. *For any $\phi:\mathbb{Z}\to\mathbb{R}$ of finite support and any $\widetilde{\lambda}>0$ we have*

$$\#\{n:\mathcal{B}^*(\phi,n)>\widetilde{\lambda}\} \leq \frac{4}{\widetilde{\lambda}}||\phi||_{\ell^1}. \tag{55}$$

Proof. Since $\mathcal{B}^*(|\phi|,n) \geq \mathcal{B}^*(\phi,n)$ and the right-hand side of (55) is unchanged if $|\phi|$ is used instead of ϕ we can suppose that $\phi \geq 0$. Using notation introduced in the proof of Lemma 4 set $\phi_{0,j,+}(n) = \phi_{t(n),0,j}(n)$. Then using $\phi \geq 0$ one can see that for a fixed t

$$\sum_{n\in[(t-1)p,tp)}\phi(n) = \sum_{n=(t-1)p}^{(t-1)p+q_j-1}\sum_{k=0}^{\widetilde{q}_j-1}\phi(n+kq_j) =$$

$$\sum_{n=(t-1)p}^{(t-1)p+q_j-1} \widetilde{q}_j \phi_{t(n),0,j}(n) = \sum_{n\in[(t-1)p,tp)} \phi_{0,j,+}(n),$$

and hence

$$||\phi_{0,j,+}||_{\ell^1} = ||\phi||_{\ell^1}. \tag{56}$$

Put

$$\mathcal{B}^*_K(\phi,n) \stackrel{\text{def}}{=} \sup_N \frac{1}{K} \sum_{j=1}^{K} \mathcal{B}(\phi,n,N,j),$$

$$\phi_{0,+}(n) = \frac{1}{K} \sum_{j=1}^{K} \phi_{0,j,+}(n),$$

and

$$\phi^*_{0,+}(n) = \sup_{N'} \frac{1}{N'} \sum_{k=0}^{N'-1} \phi_{0,+}(n+kp).$$

Next we verify that

$$\mathcal{B}^*_K(\phi,n) = \phi^*_{0,+}(n). \tag{57}$$

We use an argument similar to the one used at (51). Recall that

$$\mathcal{B}^*_K(\phi,n) = \sup_N \frac{1}{K} \sum_{j=1}^{K} \frac{1}{\nu(n,N,j)} \sum_{lq_j\in I(n,N)} \phi(n+lq_j).$$

Define $t_0(n,N)$, $t_1(n,N)$ and N' as at (32). We have

$$\frac{1}{\nu(n,N,j)} \sum_{lq_j\in I(n,N)} \phi(n+lq_j) = \tag{58}$$

$$\frac{1}{N'\widetilde{q}_j} \sum_{t=t_0(n,N)}^{t_1(n,N)} \sum_{lq_j+n\in[(t-1)\cdot p,t\cdot p)} \phi(n+lq_j) =$$

$$\frac{1}{N'} \sum_{k=0}^{N'-1} \frac{1}{\widetilde{q}_j} \sum_{lq_j+n\in[(t(n+kp)-1)\cdot p,t(n+kp)\cdot p)} \phi(n+lq_j) =$$

$$\frac{1}{N'} \sum_{k=0}^{N'-1} \phi_{t(n+kp),0,j}(n+kp) = \frac{1}{N'} \sum_{k=0}^{N'-1} \phi_{0,j,+}(n+kp).$$

Thus,

$$\frac{1}{K}\sum_{j=1}^{K}\frac{1}{\nu(n,N,j)}\sum_{lq_j\in I(n,N)}\phi(n+lq_j)=\frac{1}{N'}\sum_{k=0}^{N'-1}\frac{1}{K}\sum_{j=1}^{K}\phi_{0,j,+}(n+kp).$$

Now taking supremums in N and hence in N' we obtain (57).

By (54) and (57)

$$\mathcal{B}^*(\phi,n)\leq\sup_N\frac{2}{K}\sum_{j=1}^{K}\mathcal{B}(\phi,n,N,j)=$$

$$2\mathcal{B}_K^*(\phi,n)=2\sup_{N'>0}\frac{1}{N'}\sum_{k=0}^{N'-1}\frac{1}{K}\sum_{j=1}^{K}\phi_{0,j,+}(n+kp)=2\phi_{0,+}^*(n).$$

Hence,

$$\#\{n:\mathcal{B}^*(\phi,n)>\widetilde{\lambda}\}\leq\#\{n:\phi_{0,+}^*>\widetilde{\lambda}/2\}\leq$$

(using Lemma 5 for $n+p\mathbb{Z}$ instead of $\mathbb{Z}$ and then adding for n's)

$$2\cdot\frac{2}{\widetilde{\widetilde{\lambda}}}\sum_{n=0}^{p-1}\sum_{k\in\mathbb{Z}}\phi_{0,+}(n+kp)=\frac{4}{\widetilde{\widetilde{\lambda}}}\sum_{n\in\mathbb{Z}}\phi_{0,+}(n)=$$

(using (56))

$$\frac{4}{\widetilde{\widetilde{\lambda}}}\sum_{n\in\mathbb{Z}}\frac{1}{K}\sum_{j=1}^{K}\phi_{0,j,+}(n)=\frac{4}{\widetilde{\widetilde{\lambda}}}\frac{1}{K}\sum_{j=1}^{K}||\phi_{0,j,+}||_{\ell^1}=\frac{4}{\widetilde{\widetilde{\lambda}}}||\phi||_{\ell^1}.$$

□

In the next two sections we prove Lemma 3.

5 Part 1 of the proof of Lemma 3

Proof of Lemma 3. It is sufficient to show the lemma by assuming $f\geq 0$ and by approximating L^1 functions with simple functions we can assume that f takes finitely many values.

Suppose $\lambda>0$ is fixed. Set $X(A^*)=\{x:A^*(f,x)>\lambda\}$, and $\lambda'=\lambda/3$.

Set

$$f_{1,m}(x) = \begin{cases} f(x)/\lambda' & \text{if} \qquad f(x)/\lambda' < \overline{N}_{m-3}; \\ 0 & \text{otherwise,} \end{cases}$$

$$f_{2,m}(x) = \begin{cases} f(x)/\lambda' & \text{if} \qquad \overline{N}_{m-3} \le f(x)/\lambda' < \overline{N}_m; \\ 0 & \text{otherwise,} \end{cases}$$

and

$$f_{3,m}(x) = \begin{cases} f(x)/\lambda' & \text{if} \qquad \overline{N}_m \le f(x)/\lambda'; \\ 0 & \text{otherwise.} \end{cases}$$

It is clear that

$$\sum_{m=1}^{\infty} f_{2,m}(x) \le 3f(x)/\lambda', \tag{59}$$

and

$$f_{1,m} = 0 \text{ if } m \le 3. \tag{60}$$

If $x \in X(A^*)$ then there exists N such that $A(f, x, N) > \lambda$. Since $3f(x)/\lambda = f(x)/\lambda' = f_{1,m}(x) + f_{2,m}(x) + f_{3,m}(x)$ if $m(N)$ is chosen so that $\beta_{m(N)-1} < N \le \beta_{m(N)}$ then there exists $i \in \{1, 2, 3\}$ such that $A(f_{i,m(N)}, x, N) > 1$. Set

$$X(A, m, i) = \{x : \sup_{\beta_{m-1} < N \le \beta_m} A(f_{i,m}, x, N) > 1\},$$

and

$$X(A, m) = \{x : \sup_{\beta_{m-1} < N \le \beta_m} A(f, x, N) > \lambda\}.$$

Then

$$X(A^*) = \cup_{m=1}^{\infty} X(A, m) \subset \cup_{m=1}^{\infty} \cup_{i=1}^{3} X(A, m, i). \tag{61}$$

The functions $n(x)$ and $r(x, m)$ will be defined later. At this stage of the proof we only assume that they are measurable in x, $r(., m) : X \to \{0, 1, ..., p_m - 1\}$, $n(.) : X \to \mathbb{Z}$. Given N we let

$$I(x, m, N) = \left[-r(x, m), \left\lfloor \frac{N + r(x, m)}{p_m} \right\rfloor p_m + p_m - r(x, m) \right) \cap \mathbb{Z}, \tag{62}$$

$$\nu(x, m, N) = \# I(x, m, N), \qquad \nu(x, m, N, j) = \nu(x, m, N)/q_{j,m}.$$

From (6) it follows that

$$\frac{\nu(x, m, N, j)}{\sum_{j'=1}^{K_m} \nu(x, m, N, j')} = \frac{\nu(x, m, N)\frac{1}{q_{j,m}}}{\nu(x, m, N)Q(m)} \le \frac{2}{K_m}. \tag{63}$$

For any g defined on X we set

$$B(g,x,m,N,j) = \frac{1}{\nu(x,m,N,j)} \sum_{lq_{j,m} \in I(x,m,N)} g(T^{lq_{j,m}}x),$$

and

$$B(g,x,m,N) = \frac{\sum_{j=1}^{K_m} \nu(x,m,N,j) B(g,x,m,N,j)}{\sum_{j=1}^{K_m} \nu(x,m,N,j)}.$$

From (63) it follows that for $g \geq 0$

$$B(g,x,m,N) \leq \frac{2}{K_m} \sum_{j=1}^{K_m} B(g,x,m,N,j). \tag{64}$$

We also introduce the operator

$$A(g,x,N,m(N)) = \frac{1}{\overline{N}_0^N} \sum_{k=\overline{N}_{m(N)-1}+1}^{\overline{N}_0^N} g(T^{n_k}x),$$

and for $1 \leq m < m(N)$ the operators

$$A(g,x,N,m) = \frac{1}{\overline{N}_0^N} \sum_{k=\overline{N}_{m-1}+1}^{\overline{N}_m} g(T^{n_k}x).$$

We have

$$A(g,x,N) = \sum_{m=1}^{m(N)} A(g,x,N,m). \tag{65}$$

Next we verify that for any choice of $r(x,m)$, $n(x)$, for any $g \geq 0$, $N \in \mathbb{N}$, if $m_0 = m(N)$ then

$$2B(g,x,m_0,N) \geq A(g,x,N,m_0). \tag{66}$$

It is clear that $\nu(x,m_0,N) \leq N + 2p_{m_0}$ and by $N > \beta_{m_0-1}$, (10), and (29) we have

$$\frac{Q(m_0)\nu(x,m_0,N)}{\overline{N}_0^N} < \frac{Q(m_0)1.01 \cdot N}{\frac{3}{5}Q(m_0)N} < 2. \tag{67}$$

Now, still supposing $g \geq 0$

$$B(g, x, m_0, N) = \frac{\sum_{j=1}^{K_{m_0}} \sum_{lq_{j,m_0} \in I(x,m_0,N)} g(T^{lq_{j,m_0}} x)}{\sum_{j=1}^{K_{m_0}} \nu(x, m_0, N, j)} \geq$$

$$\frac{\sum_{n_k \in [\beta_{m_0-1}, N)} g(T^{n_k} x)}{Q(m_0)\nu(x, m_0, N)} \geq A(g, x, N, m_0) \cdot \frac{\overline{N}_0^N}{Q(m_0)\nu(x, m_0, N)},$$

that is,

$$\frac{Q(m_0)\nu(x, m_0, N)}{\overline{N}_0^N} B(g, x, m_0, N) \geq A(g, x, N, m_0),$$

and (67) implies (66).

Suppose that for an $N \in (\beta_{m_0-1}, \beta_{m_0}]$ we have $A(f_{2,m_0}, x, N) > 1$, that is, $x \in X(A, m_0, 2)$.

For $m \leq m_0$ and $\beta_{m_0-1} < N' \leq N$ set

$$X(f_{2,m_0}, N', m, \frac{1}{3}) = \{x : A(f_{2,m_0}, x, N', m) > \frac{1}{3}\},$$

and

$$X(f_{2,m_0}, N', m, +) = \{x : A(f_{2,m_0}, x, N', m) > 0\}.$$

Recall that if $f_{2,m_0}(x) \neq 0$ then

$$\overline{N}_{m_0-3} \leq f_{2,m_0}(x) < \overline{N}_{m_0}.$$

We also put

$$X(f_{2,m_0}, +) = \{x : f_{2,m_0}(x) > 0\}.$$

Then

$$\cup_{m=1}^{m_0-3} X(f_{2,m_0}, N, m, +) \subset \cup_{k=1}^{\overline{N}_{m_0-3}} T^{-n_k}(X(f_{2,m_0}, +)),$$

which implies

$$\mu(\cup_{m=1}^{m_0-3} X(f_{2,m_0}, N, m, +)) \leq \overline{N}_{m_0-3}\mu(X(f_{2,m_0}, +)) \leq \int f_{2,m_0} d\mu. \qquad (68)$$

For $m \leq m_0 - 1$ and any $N, N' \in (\beta_{m_0-1}, \beta_{m_0}]$

$$X(f_{2,m_0}, N, m, +) = X(f_{2,m_0}, N', m, +).$$

Thus from (68) we infer that

$$\mu(\{x : \sup_{\beta_{m_0-1}<N\leq\beta_{m_0}} \sum_{m=1}^{m_0-3} A(f_{2,m_0}, x, N, m) > 0\}) \leq \int f_{2,m_0} d\mu. \tag{69}$$

Next we have to estimate

$$\mu(\{x : \sup_{\beta_{m_0-1}<N\leq\beta_{m_0}} A(f_{2,m_0}, x, N, m) > \frac{1}{3}\}) \tag{70}$$

for $m = m_0 - 2$, $m_0 - 1$ and m_0.

If $m' = m_0 - 2$, or $m_0 - 1$ then for any $\beta_{m_0-1} < N \leq \beta_{m_0}$ we have $\beta_{m'} < N$ and $A(f_{2,m_0}, x, N, m') \leq A(f_{2,m_0}, x, \beta_{m'}, m')$. Hence

$$\mu(\{x : \sup_{\beta_{m_0-1}<N\leq\beta_{m_0}} A(f_{2,m_0}, x, N, m') > \frac{1}{3}\}) \leq \tag{71}$$

$$\mu(\{x : A(f_{2,m_0}, x, \beta_{m'}, m') > \frac{1}{3}\}) \leq$$

$$3 \int \frac{1}{\overline{N}_{m'}} \sum_{k=\overline{N}_{m'-1}}^{\overline{N}_{m'}} f_{2,m_0}(T^{n_k}x) d\mu(x) \leq 3 \int f_{2,m_0}(x) d\mu(x).$$

To estimate (70) when $m = m_0$ is a little more involved.

If $\int f_{2,m_0} d\mu = 0$ then we have nothing to prove. Hence, suppose

$$\epsilon_{m_0} = \int f_{2,m_0} d\mu > 0. \tag{72}$$

Later we will choose a sufficiently large κ_{m_0} and by the Kakutani-Rokhlin lemma a set E_{2,m_0} such that $E_{2,m_0}, ..., T^{\kappa_{m_0}-1}E_{2,m_0}$ are disjoint and

$$\mu(\bigcup_{k=0}^{\kappa_{m_0}-1} T^k E_{2,m_0}) > 1 - \epsilon_{m_0}. \tag{73}$$

Then $1/\mu(E_{2,m_0}) \leq 1/\kappa_{m_0}$ and we can assume that κ_{m_0} is so large that

$$(\beta_{m_0} + 3p_{m_0})\mu(E_{2,m_0}) \leq (\beta_{m_0} + 3p_{m_0})/\kappa_{m_0} < \epsilon_{m_0}. \tag{74}$$

Since f takes only finitely many values so does f_{2,m_0}. Thus, we can partition each $T^k E_{2,m_0}$ into a finite partition $\alpha_{2,m_0,k}$ so that f_{2,m_0} is constant on each

partition element. Consider $\alpha_{2,m_0} = \vee_{k=0}^{\kappa_{m_0}-1} T^{-k}\alpha_{2,m_0,k}$. If $E' \in \alpha_{2,m_0}$ then f_{2,m_0} is constant on each $T^k E'$, $k = 0, ..., \kappa_{m_0} - 1$. It is enough to deal with the E''s when $\mu(E') > 0$, and hence we suppose this.

Choose an arbitrary $x \in E'$ and set

$$\phi_{E'}(n) = f_{2,m_0}(T^n x) \text{ for } n = 0, ..., \kappa_{m_0} - 1.$$

For other n's set $\phi_{E'}(n) = 0$. If $x \notin \cup_{n=0}^{\kappa_{m_0}-1} T^n E_{2,m_0}$, then set $r(x, m_0) = 0$. If $x \in \cup_{n=0}^{\kappa_{m_0}-1} T^n E_{2,m_0}$ then there is a unique $E'(x) \in \alpha_{2,m_0}$ and $n(x)$ such that $x \in T^{n(x)} E'(x)$. In this case, set $r(x, m_0) = n(x) - \lfloor n(x)/p_{m_0} \rfloor p_{m_0}$.

Suppose $E' \in \alpha_{2,m_0}$ is fixed and $x \in \cup_{n=0}^{\kappa_{m_0}-p_{m_0}-1} T^n E'$ and $N \leq \beta_{m_0}$. Then letting $t(n(x)) = \lfloor n(x)/p_{m_0} \rfloor + 1$ we have $n(x) \in [(t(n(x))-1)p_{m_0}, t(n(x))p_{m_0})$ and if we use $p = p_{m_0}$ in (33) then taking into consideration (62)

$$I(n(x), N) = (t(n(x)) - 1)p_{m_0} - n(x) + r(x, m_0) + I(x, m_0, N) = I(x, m_0, N),$$

$$\nu(n(x), N) = \nu(x, m_0, N) \qquad \nu(n(x), N, j) = \nu(x, m_0, N, j),$$

and for $x \in \cup_{n=p_{m_0}}^{\kappa_{m_0}-\beta_{m_0}-2p_{m_0}-1} T^n E'$, with $q_j = q_{j,m_0}$ in the definition of $\mathcal{B}$,

$$\mathcal{B}(\phi_{E'}, n(x), N, j) = B(f_{2,m_0}, x, m_0, N, j),$$

and

$$\mathcal{B}(\phi_{E'}, n(x), N) = B(f_{2,m_0}, x, m_0, N).$$

By (66) for $x \in \cup_{n=p_{m_0}}^{\kappa_{m_0}-\beta_{m_0}-2p_{m_0}-1} T^n E'$

$$\sup_{\beta_{m_0-1} < N \leq \beta_{m_0}} A(f_{2,m_0}, x, N, m_0) \leq 2 \sup_{\beta_{m_0-1} < N \leq \beta_{m_0}} B(f_{2,m_0}, x, m_0, N) = \tag{75}$$

$$2 \sup_{\beta_{m_0-1} < N \leq \beta_{m_0}} \mathcal{B}(\phi_{E'}, n(x), N) \leq 2 \sup_{N>0} \mathcal{B}(\phi_{E'}, n(x), N) = 2\mathcal{B}^*(\phi_{E'}, n(x)).$$

From (55) of Lemma 7 it follows that

$$\#\{n : \mathcal{B}^*(\phi_{E'}, n) > \frac{1}{6}\} \leq 24 \sum_{n \in \mathbb{Z}} \phi_{E'}(n). \tag{76}$$

Using that $\mu(T^n E') = \mu(E')$ and the sets $T^n E'$ are disjoint for $n = 0, ..., \kappa_{m_0} - 1$, if we multiply both sides of (76) by $\mu(E')$ and take into consideration (75) then we obtain

$$\mu \left\{ x \in \bigcup_{n=p_{m_0}}^{\kappa_{m_0}-\beta_{m_0}-2p_{m_0}-1} T^n E' : \sup_{\beta_{m_0-1} < N \leq \beta_{m_0}} A(f_{2,m_0}, x, N, m_0) > \frac{1}{3} \right\} \leq \tag{77}$$

$$24 \int_{\cup_{n=0}^{\kappa_{m_0}-1} T^n E'} f_{2,m_0} d\mu.$$

Adding (77) for all $E' \in \alpha_{2,m_0}$ we have

$$\mu\left\{x \in \bigcup_{n=p_{m_0}}^{\kappa_{m_0}-\beta_{m_0}-2p_{m_0}-1} T^n E_{2,m_0} : \sup_{\beta_{m_0-1}<N\leq\beta_{m_0}} A(f_{2,m_0}, x, N, m_0) > \frac{1}{3}\right\} \leq \tag{78}$$

$$24 \int_X f_{2,m_0} d\mu.$$

This, (72), (73) and (74) imply that

$$\mu\{x : \sup_{\beta_{m_0-1}<N\leq\beta_{m_0}} A(f_{2,m_0}, x, N, m_0) > \frac{1}{3}\} \leq 26 \int f_{2,m_0} d\mu. \tag{79}$$

Now,

$$\{x : \sup_N A(f, x, N) > \lambda\} = \{x : \exists N(x),\ A(f, x, N(x)) > \lambda\} = \tag{80}$$

(we select and fix a measurable function $N(x)$)

$$= \{x : \beta_{m(N(x))-1} < N(x) \leq \beta_{m_{N(x)}},\ A(f, x, N(x)) > \lambda\} \subset$$

(by (61))

$$\bigcup_{m_0=1}^{\infty} \bigcup_{i=1}^{3} \{x : \sup_{\beta_{m_0-1}<N\leq\beta_{m_0}} A(f_{i,m_0}, x, N) > 1\}.$$

Using (65)

$$\{x : \sup_{\beta_{m_0-1}<N\leq\beta_{m_0}} A(f_{2,m_0}, x, N) > 1\} \subset$$

$$\{x : \sup_{\beta_{m_0-1}<N\leq\beta_{m_0}} \sum_{m=1}^{m_0-3} A(f_{2,m_0}, x, N, m) > 0\} \cup$$

$$\bigcup_{m'=m_0-2}^{m_0-1} \{x : \sup_{\beta_{m_0-1}<N\leq\beta_{m_0}} A(f_{2,m_0}, x, N, m') > \frac{1}{3}\} \cup$$

$$\{x : \sup_{\beta_{m_0-1}<N\leq\beta_{m_0}} A(f_{2,m_0}, x, N, m_0) > \frac{1}{3}\}.$$

This implies that by (69), (71), and (79)

$$\mu\{x : \sup_{\beta_{m_0-1}<N\le\beta_{m_0}} A(f_{2,m_0}, x, N) > 1\} \le 33 \int f_{2,m_0} d\mu,$$

and

$$\mu\left(\bigcup_{m_0=1}^{\infty}\{x : \sup_{\beta_{m_0-1}<N\le\beta_{m_0}} A(f_{2,m_0}, x, N) > 1\}\right) \le 33 \int \sum_{m_0=1}^{\infty} f_{2,m_0} d\mu \le \quad (81)$$

(by (59))

$$\frac{99\int f d\mu}{\lambda'} \le \frac{300\int f d\mu}{\lambda}.$$

The estimation for the functions of the type f_{3,m_0} is quite simple. We have

$$\bigcup_{m_0=1}^{\infty}\{x : \sup_{\beta_{m_0-1}<N\le\beta_{m_0}} A(f_{3,m_0}, x, N) > 1\} \subset \quad (82)$$

$$\bigcup_{m_0=1}^{\infty}\{x : \sup_{1\le N\le\beta_{m_0}} A(f_{3,m_0}, x, N) > 0\}.$$

Put

$$X_{3,m} = \{x : \lambda'\overline{N}_{m+1} > f(x) \ge \lambda'\overline{N}_m\}. \quad (83)$$

Observe that $A(f_{3,m_0}, x, N) = 0$ if $N \le \beta_{m_0}$ and $x \notin \cup_{k=1}^{\overline{N}_{m_0}} \cup_{m=m_0}^{\infty} T^{-n_k} X_{3,m}$. Thus,

$$\bigcup_{m_0=1}^{\infty}\{x : \sup_{1\le N\le\beta_{m_0}} A(f_{3,m_0}, x, N) > 0\} \subset \bigcup_{m_0=1}^{\infty}\bigcup_{k=1}^{\overline{N}_{m_0}}\bigcup_{m=m_0}^{\infty} T^{-n_k} X_{3,m} = \quad (84)$$

$$\bigcup_{m_0=1}^{\infty}\bigcup_{m=m_0}^{\infty}\bigcup_{k=1}^{\overline{N}_{m_0}} T^{-n_k} X_{3,m} = \bigcup_{m_0=1}^{\infty}\bigcup_{k=1}^{\overline{N}_{m_0}} T^{-n_k} X_{3,m_0}.$$

From (82), (83), and (84) it follows that

$$\mu\left(\bigcup_{m_0=1}^{\infty}\{x : \sup_{\beta_{m_0-1}<N\le\beta_{m_0}} A(f_{3,m_0}, x, N) > 1\}\right) \le \mu\left(\bigcup_{m_0=1}^{\infty}\bigcup_{k=1}^{\overline{N}_{m_0}} T^{-n_k} X_{3,m_0}\right) \le \quad (85)$$

$$\sum_{m_0=1}^{\infty} \overline{N}_{m_0}\mu(X_{3,m_0}) \le \int \frac{f(x)}{\lambda'} d\mu(x) < \frac{3\int f d\mu}{\lambda}.$$

6 Part 2 of the proof of Lemma 3

Suppose $\beta_{m_0-1} < N \leq \beta_{m_0}$.

We need to estimate $\mu\{x : \sup_{\beta_{m_0-1}<N\leq\beta_{m_0}} A(f_{1,m_0}, x, N) > 1\}$. At the beginning we argue similarly to the case $m = m_0$ when we had to obtain an estimate of the functions f_{2,m_0}, however soon this proof gets much more complicated. This is mainly due to the fact that in the earlier argument $\sum_{m_0} f_{2,m_0}$ is in L^1 while we do not have this for $\sum_{m_0} f_{1,m_0}$. To handle this problem after we have applied a Kakutani-Rokhlin tower construction we need to take advantage of the proper choice of constants K_{m_0} and of Lemma 4.

By choosing our initial parameters properly we can assume that for all $m_0 > 3$,

$$(\sum_{m=1}^{m_0-2} \overline{N}_m)\frac{\overline{N}_{m_0-3}}{\overline{N}_0^N} \leq (\sum_{m=1}^{m_0-2} \overline{N}_m)\frac{\overline{N}_{m_0-3}}{\overline{N}_{m_0-1}} < \frac{1}{100m_0}. \tag{86}$$

If $m_0 \leq 3$ then $f_{1,m_0} = 0$, hence it is enough to obtain an estimate for $m_0 > 3$.

By our assumptions and by its definition $0 \leq f_{1,m_0} < \overline{N}_{m_0-3}$ and later we will use this estimate quite often.

By (86) we have

$$\sum_{m=1}^{m_0-2} A(f_{1,m_0}, x, N, m) \leq (\sum_{m=1}^{m_0-2} \overline{N}_m)\frac{\overline{N}_{m_0-3}}{\overline{N}_0^N} < \frac{1}{100m_0}. \tag{87}$$

If $\int f_{1,m_0} d\mu = 0$ then we have nothing to prove. Hence, suppose

$$\epsilon_{1,m_0} = \min\{2^{-m_0}, 2^{-m_0}\int f_{1,m_0} d\mu\} > 0. \tag{88}$$

Later we will select a sufficiently large κ_{1,m_0} and by the Kakutani-Rokhlin lemma choose E_{1,m_0} such that $E_{1,m_0}, ..., T^{\kappa_{1,m_0}-1}E_{1,m_0}$ are disjoint and

$$\mu(\cup_{k=0}^{\kappa_{1,m_0}-1} T^k E_{1,m_0}) > 1 - \epsilon_{1,m_0}. \tag{89}$$

Then $1/\mu(E_{1,m_0}) < 1/\kappa_{1,m_0}$ and we can assume that κ_{1,m_0} is so large that

$$(\beta_{m_0} + 3p_{m_0})\mu(E_{1,m_0}) < (\beta_{m_0} + 3p_{m_0})/\kappa_{1,m_0} < \epsilon_{1,m_0}. \tag{90}$$

Since f takes only finitely many values, so does f_{1,m_0}. Thus we can divide each $T^k E_{1,m_0}$ into a finite partition $\alpha_{1,m_0,k}$ so that f_{1,m_0} is constant on each

partition element. Consider $\alpha_{1,m_0} = \vee_{k=0}^{\kappa_{1,m_0}-1} T^{-k}\alpha_{1,m_0,k}$. If $E' \in \alpha_{1,m_0}$ then f_{1,m_0} is constant on each $T^k E'$, $k = 0, ..., \kappa_{1,m_0} - 1$. It is enough to deal with the E''s when $\mu(E') > 0$, and hence we suppose this.

Choose an arbitrary $x \in E'$ and set

$$\phi_{E'}(n) = f_{1,m_0}(T^n x) \text{ for } n = 0, ..., \kappa_{1,m_0} - 1.$$

For other n's set $\phi_{E'}(n) = 0$. If $x \notin \cup_{n=0}^{\kappa_{1,m_0}-1} T^n E_{1,m_0}$, or $m > m_0$ then set $r(x, m) = 0$. If $x \in \cup_{n=0}^{\kappa_{1,m_0}-1} T^n E_{1,m_0}$ then there is a unique $E'(x) \in \alpha_{1,m_0}$ and $n(x)$ such that $x \in T^{n(x)} E'(x)$, in this case for $m' \leq m_0$ set $r(x, m') = n(x) - \lfloor n(x)/p_{m'} \rfloor p_{m'}$, $t(n(x), m') = \lfloor \frac{n(x)}{p_{m'}} \rfloor + 1$. This means that $n(x) \in [(t(n(x), m') - 1)p_{m'}, t(n(x), m')p_{m'})$ and if we use $p = p_{m'}$ in (33) then by using (62) we have

$$I(n(x), N) = (t(n(x), m') - 1)p_{m'} - n(x) + r(x, m') + I(x, m', N) = I(x, m', N).$$

Still using $p = p_{m'}$ and $q_j = q_{j,m'}$ in (33) and (34) set

$$\nu(x, m', N) = \nu(n(x), N), \ \nu(x, m', N, j) = \nu(n(x), N, j).$$

We also put

$$t_0(x, m', N) = t(n(x), m'), \ t_1(x, m', N) = t(n(x), m') + \nu(x, m', N)/p_{m'}.$$

For $x \in \cup_{n=p_{m'}}^{\kappa_{1,m_0}-p_{m'}-1} T^n E'$ set

$$\overline{f}_{1,m_0}(x, m') = \frac{1}{p_{m'}} \sum_{k \in I(x,m',0)} f_{1,m_0}(T^k x) =$$

$$\frac{1}{p_{m'}} \sum_{k=(t(n(x),m')-1)p_{m'}}^{t(n(x),m')p_{m'}-1} \phi_{E'}(k) \stackrel{\text{def}}{=} \overline{\phi}_{E',m'}(n(x)).$$

If $x \notin \cup_{n=p_{m'}}^{\kappa_{1,m_0}-p_{m'}-1} T^n E'$ set $\overline{f}_{1,m_0}(x, m') = 0$.

For $x \in \cup_{n=p_{m'}}^{\kappa_{1,m_0}-\beta_{m_0}-p_{m_0}-1} T^n E'$ and $0 \leq N \leq \beta_{m_0}$ set

$$B_0(f_{1,m_0}, x, m', N, j) = \frac{1}{\nu(x, m', N, j)} \cdot$$

$$\cdot \sum_{t=t_0(x,m',N)}^{t_1(x,m',N)} \left| \sum_{lq_{j,m'}+n(x) \in [(t-1)p_{m'}, tp_{m'})} f_{1,m_0}(T^{lq_{j,m'}} x) - \overline{f}_{1,m_0}(T^{lq_{j,m'}} x, m') \right|,$$

and $B_0(f_{1,m_0},x,m',N) = \frac{\sum_{j=1}^{K_{m'}} \nu(x,m',N,j)B_0(f_{1,m_0},x,m',N,j)}{\sum_{j=1}^{K_{m'}} \nu(x,m',N,j)}$. Observe that (for $N \leq \beta_{m_0}$)

$$B_0(f_{1,m_0},x,m',N,j) = \mathcal{B}_0(\phi_{E'},n(x),N,j),$$

and hence

$$B_0(f_{1,m_0},x,m',N) = \mathcal{B}_0(\phi_{E'},n(x),N),$$

provided $p = p_{m'}$, $q_j = q_{j,m'}$, $j = 1,...,K_{m'}$ are used in the definition of $\mathcal{B}_0$. To emphasize this dependence on m' we will use the notation

$$\mathcal{B}_0(\phi_{E'},n(x),m',N,j) = \mathcal{B}_0(\phi_{E'},n(x),N,j),$$

and

$$\mathcal{B}_0(\phi_{E'},n(x),m',N) = \mathcal{B}_0(\phi_{E'},n(x),N),$$

when the above choice of parameters is used.

Set

$$I_1(x,m_0,N) = \mathbb{Z}\cap \left[\left\lfloor \frac{\beta_{m_0-1}+r(x,m_0)}{p_{m_0}} \right\rfloor p_{m_0} - r(x,m_0), \left\lfloor \frac{N+r(x,m_0)}{p_{m_0}} \right\rfloor p_{m_0} + p_{m_0} - r(x,m_0)\right),$$

for $1 \leq m < m_0$ set

$$I_1(x,m,N) = \mathbb{Z}\cap \left[\left\lfloor \frac{\beta_{m-1}+r(x,m)}{p_m} \right\rfloor p_m - r(x,m), \left\lfloor \frac{\beta_m+r(x,m)}{p_m} \right\rfloor p_m + p_m - r(x,m)\right).$$

We also put for $1 \leq m \leq m_0$

$$\nu_1(x,m,N) = \#I_1(x,m,N), \qquad \nu_1(x,m,N,j) = \frac{\nu_1(x,m,N)}{q_{j,m}}.$$

Next we need some estimates. We also use the notation introduced in the end of Section 3. Clearly, for $m < m_0$

$$P_m p_m \leq \nu_1(x,m,N) \leq (P_m+2)p_m, \tag{91}$$

and

$$P_{m_0,N}p_{m_0} \leq \nu_1(x,m_0,N) \leq (P_{m_0,N}+2)p_{m_0}. \tag{92}$$

$$N - \beta_{m_0-1} \leq \nu_1(x, m_0, N) \leq (P_{m_0,N} + 2)p_{m_0} < N - \beta_{m_0-1} + 2p_{m_0}. \quad (93)$$

By (10) and (20)

$$\beta_{m_0-1}(1-\gamma_\beta) < \beta_{m_0-1} - \beta_{m_0-2} < \nu_1(x, m_0-1, N) \leq \quad (94)$$

$$(P_{m_0-1} + 2)p_{m_0-1} < \beta_{m_0-1} - \beta_{m_0-2} + 2p_{m_0-1} < \beta_{m_0-1},$$

From (27) and (28) it follows that

$$|\overline{N}^N_{\beta_{m_0-1}} - Q(m_0)(N - \beta_{m_0-1})| < \gamma_{m_0} Q(m_0)(N - \beta_{m_0-1}) + p_{m_0} Q(m_0). \quad (95)$$

On the other hand, by the definition of $I_1(x, m_0, N)$ and $\nu_1(x, m_0, N)$

$$|Q(m_0)\nu_1(x, m_0, N) - Q(m_0)(N - \beta_{m_0-1})| < 2p_{m_0} Q(m_0). \quad (96)$$

Hence,

$$|\overline{N}^N_{\beta_{m_0-1}} - Q(m_0)\nu_1(x, m_0, N)| < \gamma_{m_0} Q(m_0)(N - \beta_{m_0-1}) + 3p_{m_0} Q(m_0). \quad (97)$$

By (27) and (92)

$$\overline{N}^N_{\beta_{m_0-1}} > (1-\gamma_{m_0}) P_{m_0,N} p_{m_0} Q(m_0) > \quad (98)$$

$$(1-\gamma_{m_0})(\nu_1(x, m_0, N) - 2p_{m_0}) Q(m_0).$$

From (28) and (92) it follows that

$$\overline{N}^N_{\beta_{m_0-1}} < (\nu_1(x, m_0, N) + p_{m_0}) Q(m_0). \quad (99)$$

Using (98) and (99) we infer

$$|\overline{N}^N_{\beta_{m_0-1}} - \nu_1(x, m_0, N) Q(m_0)| < \gamma_{m_0} \nu_1(x, m_0, N) Q(m_0) + 2p_{m_0} Q(m_0). \quad (100)$$

By (24) and (94)

$$\overline{N}^{\beta_{m_0-1}}_{\beta_{m_0-2}} < (\beta_{m_0-1} - \beta_{m_0-2} + p_{m_0-1}) Q(m_0 - 1) < \quad (101)$$

$$(\nu_1(x, m_0-1, N) + p_{m_0-1}) Q(m_0 - 1).$$

On the other hand, by (23) and (94)

$$\overline{N}^{\beta_{m_0-1}}_{\beta_{m_0-2}} > (1-\gamma_{m_0-1})(\nu_1(x, m_0-1, N) - 3p_{m_0-1}) Q(m_0 - 1). \quad (102)$$

From (101) and (102) we infer

$$|\overline{N}_{\beta_{m_0-2}}^{\beta_{m_0-1}} - \nu_1(x, m_0 - 1, N)Q(m_0 - 1)| < \tag{103}$$

$$\gamma_{m_0-1}\nu_1(x, m_0 - 1, N)Q(m_0 - 1) + 3p_{m_0-1}Q(m_0 - 1).$$

For $1 \leq m \leq m_0$ set

$$S_1(f_{1,m_0}, x, m, N, j) = \sum_{lq_{j,m} \in I_1(x,m,N)} f_{1,m_0}(T^{lq_{j,m}}x),$$

$$\overline{S}_1(f_{1,m_0}, x, m, N, j) = \sum_{lq_{j,m} \in I_1(x,m,N)} \overline{f}_{1,m_0}(T^{lq_{j,m}}x, m),$$

$$S_1(f_{1,m_0}, x, m, N) = \sum_{j=1}^{K_m} S_1(f_{1,m_0}, x, m, N, j),$$

and

$$\overline{S}_1(f_{1,m_0}, x, m, N) = \sum_{j=1}^{K_m} \overline{S}_1(f_{1,m_0}, x, m, N, j).$$

Until the end of the proof of this lemma we assume that $m' = m_0 - 1$, or m_0.

Recall that in any subinterval of length $p_{m'}$ belonging to $[\beta_{m'-1}, \beta_{m'})$ the sets $\Lambda_{j,m',0}$ have $\widetilde{q}_{j,m'} = p_{m'}/q_{j,m'}$ many elements. From $\Lambda_{j,m',0}$ during the definition of $\Lambda_{m'}$ (see (15) and the paragraph above it) less than

$$\sum_{j' \neq j} \widetilde{q}_{j,m'} \frac{1}{q_{j',m'}} 2(d_{m'} + 1) \tag{104}$$

many elements are deleted. The intervals $[\beta_{m_0-2}, \beta_{m_0-1}) \cap \mathbb{Z}$ and $I_1(x, m_0 - 1, N)$ are roughly the same, apart from two intervals of cardinality no more than p_{m_0-1} at the beginning and in the end, to state this more precisely

$$\#(([\beta_{m_0-2}, \beta_{m_0-1}) \cap \mathbb{Z})\Delta I_1(x, m_0 - 1, N)) \leq 2p_{m_0-1}, \tag{105}$$

where Δ stands for the symmetric difference. Similarly,

$$\#(([\beta_{m_0-1}, N) \cap \mathbb{Z})\Delta I_1(x, m_0, N)) \leq 2p_{m_0}, \tag{106}$$

or, by changing by one element at the beginning and in the end

$$\#((\beta_{m_0-1}, N] \cap \mathbb{Z}) \Delta I_1(x, m_0, N) < 3p_{m_0}. \tag{107}$$

If $\mathcal{N}(m')$ denotes the total number of grid intervals of length $p_{m'}$ which are shifted by $-r(x, m')$ and are belonging to $I_1(x, m', N)$ then

$$\mathcal{N}(m') = \frac{\nu_1(x, m', N, j)}{\widetilde{q}_{j,m'}}, \text{ for any } j = 1, ..., K_{m'}. \tag{108}$$

Next we verify that by our choice of the initial parameters we have

$$\sum_{j=1}^{K_{m'}} \nu_1(x, m', N, j) < \sum_{j=1}^{K_{m'}} \nu(x, m', \min\{N, \beta_{m'}\}, j) < 2\overline{N}_0^N. \tag{109}$$

holds.

Observe that $\min\{N, \beta_{m'}\}$ equals N when $m' = m_0$ and equals β_{m_0-1} if $m' = m_0 - 1$.

Since $N > \beta_{m_0-1}$ and $m' \in \{m_0, m_0 - 1\}$ by (10) we have

$$N + p_{m'} < 1.01 \cdot N \text{ and } \beta_{m_0-1} + p_{m'} < 1.01 \cdot \beta_{m_0-1}. \tag{110}$$

By (19) and (29) we have

$$\overline{N}_0^N \geq \frac{3}{4} \cdot \frac{98}{100}(\beta_{m_0-1} - \beta_{m_0-2})Q(m_0 - 1) > \frac{3}{4} \cdot \frac{98}{100} \cdot \frac{999}{1000}\beta_{m_0-1}Q(m_0 - 1)$$

and

$$\overline{N}_0^N \geq \frac{3}{5}Q(m_0)N.$$

By the definition of $I(x, m', \min\{N, \beta_{m'}\})$ and (110) we have

$$\nu(x, m', \min\{N, \beta_{m'}\}) < \min\{N, \beta_{m'}\} + p_{m'} < 1.01 \cdot \min\{N, \beta_{m'}\}.$$

Therefore

$$\sum_{j=1}^{K_{m'}} \nu(x, m', \min\{N, \beta_{m'}\}, j) = \nu(x, m', \min\{N, \beta_{m'}\})Q(m') \leq \tag{111}$$

$$1.01 \min\{N, \beta_{m'}\}Q(m') < 2\overline{N}_0^N.$$

We also make the following assumption about our initial parameters:

$$\frac{\overline{N}_{m_0-3}}{\overline{\overline{N}}_{m_0-1}} 3p_{m_0} < \frac{1}{200m_0} \text{ for any } m_0 > 3. \tag{112}$$

From (112) it follows that if $\beta_{m_0-1} < N \leq \beta_{m_0}$ then

$$\frac{\overline{N}_{m_0-3}}{\overline{N}_0^N} 3p_{m_0-1} < \frac{\overline{N}_{m_0-3}}{\overline{N}_{m_0-1}} 3p_{m_0} < \frac{1}{200m_0}. \tag{113}$$

Using (8) and (104-113) for $m' = m_0 - 1$, or m_0 we have

$$\left| \frac{S_1(f_{1,m_0}, x, m', N)}{\overline{N}_0^N} - A(f_{1,m_0}, x, N, m') \right| \leq \tag{114}$$

$$\frac{\overline{N}_{m_0-3}}{\overline{N}_0^N} \left| 3p_{m'} + \sum_{j=1}^{K_{m'}} \nu_1(x, m', N, j) \sum_{j' \neq j} \frac{2(d_{m'}+1)}{\min\{q_{j',m'}\}} \right| \leq$$

$$\frac{\overline{N}_{m_0-3}}{\overline{N}_0^N} \left| 3p_{m'} + \sum_{j=1}^{K_{m'}} \nu_1(x, m', N, j) K_{m'} \frac{2(d_{m'}+1)}{\min\{q_{j',m'}\}} \right| < \frac{1}{100m_0}.$$

Next we estimate

$$\left| \frac{S_1(f_{1,m_0}, x, m', N) - \overline{S}_1(f_{1,m_0}, x, m', N)}{\overline{N}_0^N} \right| = \tag{115}$$

$$\left| \frac{\sum_{j=1}^{K_{m'}} \sum_{lq_{j,m'} \in I_1(x,m',N)} f_{1,m_0}(T^{lq_{j,m'}} x) - \overline{f}_{1,m_0}(T^{lq_{j,m'}} x, m')}{\overline{N}_0^N} \right| \leq$$

(using (111), $I_1(x, m', N) \subset I(x, m', N)$ and the triangle inequality)

$$\frac{2}{\sum_{j=1}^{K_{m'}} \nu(x, m', \min\{N, \beta_{m'}\}, j)} \cdot$$

$$\cdot \sum_{j=1}^{K_{m'}} \sum_{t=t_0(x,m',\min\{N,\beta_{m'}\})}^{t_1(x,m',\min\{N,\beta_{m'}\})} \left| \sum_{lq_{j,m'}+n(x) \in [(t-1)p_{m'}, tp_{m'})} f_{1,m_0}(T^{lq_{j,m'}} x) - \right.$$

$$\left. \overline{f}_{1,m_0}(T^{lq_{j,m'}} x, m') \right| =$$

$$2B_0(f_{1,m_0},x,m',\min\{N,\beta_{m'}\}) \le 2 \max_{\beta_{m'-1}<N'\le\beta_{m'}} B_0(f_{1,m_0},x,m',N').$$

Since $\overline{f}_{1,m_0}(T^{lq_{j,m'}}x,m')$ equals $\overline{f}_{1,m_0}(T^{t'p_{m'}}x,m')$ when $lq_{j,m'} \in [t'p_{m'},(t'+1)p_{m'})$ and for each t' and j there are $\widetilde{q}_{j,m'}$ many such $lq_{j,m'}$'s we have

$$\overline{S}_1(f_{1,m_0},x,m',N) = \sum_{j=1}^{K_{m'}} \widetilde{q}_{j,m'} \sum_{t'p_{m'}\in I_1(x,m',N)} \overline{f}_{1,m_0}(T^{t'p_{m'}}x,m') = \tag{116}$$

$$\sum_{j=1}^{K_{m'}} \frac{p_{m'}}{q_{j,m'}} \left(\frac{1}{p_{m'}} \sum_{k\in I_1(x,m',N)} f_{1,m_0}(T^k x) \right) = Q(m') \sum_{k\in I_1(x,m',N)} f_{1,m_0}(T^k x).$$

This implies

$$|\overline{S}_1(f_{1,m_0},x,m',N)| < \overline{N}_{m_0-3} Q(m') \nu_1(x,m',N). \tag{117}$$

We also have

$$\overline{N}_0^N = \left(\sum_{m=1}^{m_0-2} \overline{N}_{\beta_{m-1}}^{\beta_m} \right) + \overline{N}_{\beta_{m_0-2}}^{\beta_{m_0-1}} + \overline{N}_{\beta_{m_0-1}}^{N}, \tag{118}$$

and the initial parameters can be chosen so that we can estimate the sum on the right-hand side by

$$\sum_{m=1}^{m_0-2} \overline{N}_{\beta_{m-1}}^{\beta_m} < \frac{1}{100m_0} \overline{N}_{\beta_{m_0-2}}^{\beta_{m_0-1}}. \tag{119}$$

By (94) and a suitable assumption about our initial parameters

$$\nu_1(x,m_0-1,N) > (1-\gamma_\beta)\beta_{m_0-1} \text{ and} \tag{120}$$

$$\frac{3p_{m_0-1}}{\nu_1(x,m_0-1,N)} < \frac{3p_{m_0-1}}{(1-\gamma_\beta)\beta_{m_0-1}} < \frac{1}{200m_0\overline{N}_{m_0-3}}.$$

From (9), (103) and (120) it follows that

$$\left| \frac{1}{\overline{N}_{\beta_{m_0-2}}^{\beta_{m_0-1}}} - \frac{1}{\nu_1(x,m_0-1,N)Q(m_0-1)} \right| = \tag{121}$$

$$\left| \frac{\nu_1(x,m_0-1,N)Q(m_0-1) - \overline{N}_{\beta_{m_0-2}}^{\beta_{m_0-1}}}{\nu_1(x,m_0-1,N)Q(m_0-1)\overline{N}_{\beta_{m_0-2}}^{\beta_{m_0-1}}} \right| <$$

$$\frac{1}{\overline{N}_{\beta_{m_0-2}}^{\beta_{m_0-1}}}\left(\gamma_{m_0-1}+\frac{3p_{m_0-1}}{\nu_1(x,m_0-1,N)}\right)<\frac{1}{\overline{N}_{\beta_{m_0-2}}^{\beta_{m_0-1}}}\cdot\frac{1}{100\cdot m_0\overline{N}_{m_0-3}}.$$

Using (116) and (121)

$$\left|\frac{\overline{S}_1(f_{1,m_0},x,m_0-1,N)}{\overline{N}_{\beta_{m_0-2}}^{\beta_{m_0-1}}}-\frac{1}{\nu_1(x,m_0-1,N)}\sum_{k\in I_1(x,m_0-1,N)}f_{1,m_0}(T^kx)\right|< \tag{122}$$

$$\overline{S}_1(f_{1,m_0},x,m_0-1,N)\frac{1}{\overline{N}_{m_0-3}100m_0\overline{N}_{\beta_{m_0-2}}^{\beta_{m_0-1}}}<$$

(By (10), (20), (102) and (117))

$$\frac{\overline{N}_{m_0-3}Q(m_0-1)\nu_1(x,m_0-1,N)}{\overline{N}_{m_0-3}100m_0\overline{N}_{\beta_{m_0-2}}^{\beta_{m_0-1}}}<\frac{2}{100m_0}.$$

To obtain an estimate similar to (122) for m_0 instead of m_0-1. we separate two cases.
CASE 1 holds if $N-\beta_{m_0-1}\geq 10^4(m_0+1)\overline{N}_{m_0-2}p_{m_0}$, and
CASE 2 holds when $0\leq N-\beta_{m_0-1}<10^4(m_0+1)\overline{N}_{m_0-2}p_{m_0}$.

If CASE 1 holds by (116) we have

$$\left|\frac{\overline{S}_1(f_{1,m_0},x,m_0,N)}{\overline{N}_{\beta_{m_0-1}}^{N}}-\frac{\sum_{k\in I_1(x,m_0,N)}f_{1,m_0}(T^kx)}{\nu_1(x,m_0,N)}\right|= \tag{123}$$

$$\overline{S}_1(f_{1,m_0},x,m_0,N)\left|\frac{1}{\overline{N}_{\beta_{m_0-1}}^{N}}-\frac{1}{\nu_1(x,m_0,N)Q(m_0)}\right|\leq$$

(using (117))

$$Q(m_0)\nu_1(x,m_0,N)\overline{N}_{m_0-3}\frac{|\overline{N}_{\beta_{m_0-1}}^{N}-Q(m_0)\nu_1(x,m_0,N)|}{\overline{N}_{\beta_{m_0-1}}^{N}Q(m_0)\nu_1(x,m_0,N)}=$$

$$\frac{\overline{N}_{m_0-3}|\overline{N}_{\beta_{m_0-1}}^{N}-Q(m_0)\nu_1(x,m_0,N)|}{\overline{N}_{\beta_{m_0-1}}^{N}}<$$

(using (27) and (97))

$$\frac{\overline{N}_{m_0-3}(\gamma_{m_0}Q(m_0)(N-\beta_{m_0-1})+3p_{m_0}Q(m_0))}{(1-\gamma_{m_0})(N-\beta_{m_0-1}-p_{m_0})Q(m_0)}<$$

(using (9) and that for CASE 1 we have $\gamma_{m_0}(N-\beta_{m_0-1}) > 3p_{m_0}$, $N-\beta_{m_0-1} > 2p_{m_0}$ and $\gamma_{m_0} < 1/2$)

$$\frac{\overline{N}_{m_0-3}4\gamma_{m_0}}{1-\gamma_{m_0}} < 8\overline{N}_{m_0-3}\gamma_{m_0} < \frac{1}{100m_0}.$$

If CASE 2 holds then

$$|A(f_{1,m_0}, x, N, m_0)| = \frac{1}{\overline{N}_0^N} \sum_{n_k \in [\beta_{m_0-1}, N)} f_{1,m_0}(T^{n_k}x) < \tag{124}$$

$$\frac{1}{\overline{N}_0^N}(N-\beta_{m_0-1})\overline{N}_{m_0-3} < \frac{10^4(m_0+1)\overline{N}_{m_0-2}p_{m_0}\overline{N}_{m_0-3}}{\overline{N}_{m_0-1}} < \frac{1}{1000m_0},$$

where the last inequality holds if a suitable assumption is made about our initial parameters.

For both CASEs we also have

$$\left|\frac{1}{\nu_1(x,m_0-1,N)} \sum_{k\in I_1(x,m_0-1,N)} f_{1,m_0}(T^k x) - \frac{1}{\beta_{m_0-1}} \sum_{k=1}^{\beta_{m_0-1}} f_{1,m_0}(T^k x)\right| \leq \tag{125}$$

$$\left|\frac{1}{\nu_1(x,m_0-1,N)} - \frac{1}{\beta_{m_0-1}}\right| \cdot \sum_{k\in I_1(x,m_0-1,N)} f_{1,m_0}(T^k x) +$$

$$\frac{1}{\beta_{m_0-1}}\left|\sum_{k=1}^{\beta_{m_0-1}} f_{1,m_0}(T^k x) - \sum_{k\in I_1(x,m_0-1,N)} f_{1,m_0}(T^k x)\right| \leq$$

(using (105))

$$\frac{|\nu_1(x,m_0-1,N)-\beta_{m_0-1}|}{\nu_1(x,m_0-1,N)\beta_{m_0-1}} \sum_{k\in I_1(x,m_0-1,N)} f_{1,m_0}(T^k x) + \frac{(\beta_{m_0-2}+2p_{m_0-1})\overline{N}_{m_0-3}}{\beta_{m_0-1}} \leq$$

(using that (105) implies $|\nu_1(x,m_0-1,N)-\beta_{m_0-1}| \leq \beta_{m_0-2}+2p_{m_0-1}$)

$$\frac{2(\beta_{m_0-2}+2p_{m_0-1})\overline{N}_{m_0-3}}{\beta_{m_0-1}} < \frac{1}{100m_0},$$

where at the last inequality we again made an assumption about our initial parameters, especially we used that $p_{m_0-1} < \beta_{m_0-2}$ can be supposed to be much less than β_{m_0-1}.

Next observe that by (107)

$$\left|\sum_{k\in I_1(x,m_0,N)} f_{1,m_0}(T^kx)-\left(\sum_{k=1}^{N} f_{1,m_0}(T^kx)-\sum_{k=1}^{\beta_{m_0-1}} f_{1,m_0}(T^kx)\right)\right| < 3p_{m_0}\overline{N}_{m_0-3}. \tag{126}$$

It is also clear from (106) that

$$|\nu_1(x,m_0,N)-(N-\beta_{m_0-1})| \le 2p_{m_0}, \tag{127}$$

furthermore $p_{m_0} > 1$, $N-\beta_{m_0-1} \ge 1$ and (93) imply

$$\nu_1(x,m_0,N) < 3p_{m_0}(N-\beta_{m_0-1}). \tag{128}$$

By (127)

$$\left|\frac{1}{\nu_1(x,m_0,N)}-\frac{1}{N-\beta_{m_0-1}}\right| = \frac{|\nu_1(x,m_0,N)-(N-\beta_{m_0-1})|}{\nu_1(x,m_0,N)(N-\beta_{m_0-1})} \le \tag{129}$$

$$\frac{2p_{m_0}}{\nu_1(x,m_0,N)(N-\beta_{m_0-1})}.$$

Hence,

$$\left|\frac{1}{\nu_1(x,m_0,N)}\sum_{k\in I_1(x,m_0,N)} f_{1,m_0}(T^kx)- \right. \tag{130}$$

$$\left.\frac{\sum_{k=1}^{N} f_{1,m_0}(T^kx)-\sum_{k=1}^{\beta_{m_0-1}} f_{1,m_0}(T^kx)}{N-\beta_{m_0-1}}\right| <$$

(by using (126))

$$\frac{1}{\nu_1(x,m_0,N)}3p_{m_0}\overline{N}_{m_0-3}+\left|\frac{1}{\nu_1(x,m_0,N)}-\frac{1}{N-\beta_{m_0-1}}\right|(N-\beta_{m_0-1})\overline{N}_{m_0-3} \le$$

(by (93), (127) and (129))

$$\frac{3p_{m_0}\overline{N}_{m_0-3}}{(N-\beta_{m_0-1})}+\frac{2p_{m_0}}{\nu_1(x,m_0,N)}\overline{N}_{m_0-3} \le$$

$$\frac{5p_{m_0}\overline{N}_{m_0-3}}{N-\beta_{m_0-1}} \stackrel{\text{def}}{=} \mathcal{E}.$$

If CASE 1 holds, that is, $N - \beta_{m_0-1} \geq 10^4 p_{m_0} \overline{N}_{m_0-2}(m_0+1)$ then

$$\mathcal{E} < \frac{1}{100 m_0}. \tag{131}$$

Otherwise, if CASE 2 holds then

$$0 < \overline{N}_0^N - \overline{N}_{\beta_{m_0-2}}^{\beta_{m_0-1}} < \beta_{m_0-2} + 10^4 p_{m_0} \overline{N}_{m_0-2}(m_0+1). \tag{132}$$

By (103)

$$\left| \frac{\overline{N}_{\beta_{m_0-2}}^{\beta_{m_0-1}}}{Q(m_0-1)\nu_1(x, m_0-1, N)} - 1 \right| < \gamma_{m_0-1} + \frac{3p_{m_0-1}}{\nu_1(x, m_0-1, N)} < \tag{133}$$

(using (9), (10) and (94))

$$< \frac{1}{2000} + \frac{3p_{m_0-1}}{\beta_{m_0-1} - \beta_{m_0-2}} < \frac{1}{1000}.$$

Hence,

$$\frac{Q(m_0-1)\nu_1(x, m_0-1, N)}{\overline{N}_{\beta_{m_0-2}}^{\beta_{m_0-1}}} < 2. \tag{134}$$

By (117)

$$|\overline{S}_1(f_{1,m_0}, x, m_0-1, N)| < \overline{N}_{m_0-3} Q(m_0-1)\nu_1(x, m_0-1, N),$$

therefore,

$$\left| \frac{\overline{S}_1(f_{1,m_0}, x, m_0-1, N)}{\overline{N}_0^N} - \frac{\overline{S}_1(f_{1,m_0}, x, m_0-1, N)}{\overline{N}_{\beta_{m_0-2}}^{\beta_{m_0-1}}} \right| < \tag{135}$$

$$\overline{N}_{m_0-3} Q(m_0-1)\nu_1(x, m_0-1, N) \frac{|\overline{N}_0^N - \overline{N}_{\beta_{m_0-2}}^{\beta_{m_0-1}}|}{\overline{N}_0^N \overline{N}_{\beta_{m_0-2}}^{\beta_{m_0-1}}} <$$

(by (132) and (134))

$$\frac{\overline{N}_{m_0-3} 2(\beta_{m_0-2} + 10^4 p_{m_0} \overline{N}_{m_0-2}(m_0+1))}{\overline{N}_0^N} \leq$$

$$\frac{\overline{N}_{m_0-3}2(\beta_{m_0-2}+10^4 p_{m_0}\overline{N}_{m_0-2}(m_0+1))}{\overline{N}_{m_0-1}}<\frac{1}{200m_0}$$

if a suitable assumption is made about our initial parameters.

Furthermore,

$$\left|\frac{1}{N}\sum_{k=1}^{N} f_{1,m_0}(T^k x)-\frac{1}{\beta_{m_0-1}}\sum_{k=1}^{\beta_{m_0-1}} f_{1,m_0}(T^k x)\right|< \tag{136}$$

$$\left|\frac{1}{N}-\frac{1}{\beta_{m_0-1}}\right|\sum_{k=1}^{\beta_{m_0-1}} f_{1,m_0}(T^k x)+\frac{1}{N}\sum_{k=\beta_{m_0-1}+1}^{N} f_{1,m_0}(T^k x)<$$

$$\frac{N-\beta_{m_0-1}}{N}\left(\frac{1}{\beta_{m_0-1}}\sum_{k=1}^{\beta_{m_0-1}} f_{1,m_0}(T^k x)\right)+\frac{1}{N}\overline{N}_{m_0-3}(N-\beta_{m_0-1})\leq$$

(recalling that CASE 2 holds)

$$2\cdot\frac{N-\beta_{m_0-1}}{N}\overline{N}_{m_0-3}<2\cdot 10^4 p_{m_0}\overline{N}_{m_0-2}(m_0+1)\frac{1}{N}\overline{N}_{m_0-3}<$$

$$2\cdot 10^4 p_{m_0}\overline{N}_{m_0-2}(m_0+1)\frac{1}{\beta_{m_0-1}}\overline{N}_{m_0-3}<\frac{1}{200m_0},$$

if proper assumptions are made about our initial parameters.

To make easier to follow estimate (137) in an abbreviated form we recall that

by (114), $|S_1/\overline{N}_0^N - A|<1/(100m_0)$,
by (115), $|(S_1-\overline{S}_1)/\overline{N}_0^N|\leq 2\max B_0$,
by (135), $|(\overline{S}_1/\overline{N}_0^N)-(\overline{S}_1/\overline{N}_{\beta_{m_0-2}}^{\beta_{m_0-1}})|<1/(200m_0)$,
by (122), $|(\overline{S}_1/\overline{N}_{\beta_{m_0-2}}^{\beta_{m_0-1}})-(1/\nu_1)\sum_{I_1} f_{1,m_0}|<2/(100m_0)$,
by (125), $|((1/\nu_1)\sum_{I_1} f_{1,m_0})-((1/\beta_{m_0-1})\sum_1^{\beta_{m_0-1}} f_{1,m_0})|<1/(100m_0)$
and by (136), $|((1/N)\sum_1^N f_{1,m_0})-((1/\beta_{m_0-1})\sum_1^{\beta_{m_0-1}} f_{1,m_0})|<1/(200m_0)$.

Thus in CASE 2 by (87), (114), (115), (118), (122), (124), (125), (135) and (136)

$$\left|A(f_{1,m_0},x,N)-\frac{1}{N}\sum_{k=1}^{N} f_{1,m_0}(T^k x)\right|<\left|\sum_{m=1}^{m_0-2}A(f_{1,m_0},x,N,m)\right|+ \tag{137}$$

$$\left|A(f_{1,m_0},x,N,m_0-1)-\frac{1}{N}\sum_{k=1}^{N}f_{1,m_0}(T^kx)\right|+|A(f_{1,m_0},x,N,m_0)|<$$

$$\frac{6}{100m_0}+2\max_{\beta_{m_0-2}<N'\le\beta_{m_0-1}}B_0(f_{1,m_0},x,m_0-1,N')+\frac{1}{1000m_0}<$$

$$\frac{1}{10m_0}+2\max_{\beta_{m_0-2}<N'\le\beta_{m_0-1}}B_0(f_{1,m_0},x,m_0-1,N').$$

Next we need similar type estimates for CASE 1.

By the assumption for CASE 1, $N-\beta_{m_0-1}\ge 10^4(m_0+1)\overline{N}_{m_0-2}p_{m_0}$, moreover by (28), $\overline{N}^N_{\beta_{m_0-1}}<(N-\beta_{m_0-1}+p_{m_0})Q(m_0)$, and by (29), $\overline{N}^N_0>\frac{3}{5}NQ(m_0)$. Thus

$$\frac{\overline{N}^N_{\beta_{m_0-1}}}{\overline{N}^N_0}<\frac{5}{3}\frac{N-\beta_{m_0-1}+p_{m_0}}{N}<\frac{5}{3}\frac{N-\beta_{m_0-1}}{N}\left(1+\frac{p_{m_0}}{N-\beta_{m_0-1}}\right)\le \tag{138}$$

$$\frac{5}{3}\frac{N-\beta_{m_0-1}}{N}\left(1+\frac{1}{10^4(m_0+1)\overline{N}_{m_0-2}}\right)<\frac{2(N-\beta_{m_0-1})}{N}<2.$$

If CASE 1 holds using (114), (115), (118), (122) and (125) (see the list of abbreviated estimates before (137) as well)

$$\left|A(f_{1,m_0},x,N,m_0-1)-\frac{\overline{N}^{\beta_{m_0-1}}_{\beta_{m_0-2}}}{\overline{N}^N_0}\frac{1}{\beta_{m_0-1}}\sum_{k=1}^{\beta_{m_0-1}}f_{1,m_0}(T^kx)\right|\le \tag{139}$$

$$2\max_{\beta_{m_0-2}<N'\le\beta_{m_0-1}}B_0(f_{1,m_0},x,m_0-1,N')+\frac{1}{10m_0}.$$

In addition to the list of abbreviated estimates given before (137) we also recall that

by (123) we have $|(\overline{S}_1/\overline{N}^N_{\beta_{m_0-1}})-(1/\nu_1)(\sum_{k\in I_1}f_{1,m_0})|<1/(100m_0)$,

moreover by (130) and (131) we have

$|(1/\nu_1)(\sum_{k\in I_1}f_{1,m_0})-(\sum_{k=1}^{N}f_{1,m_0}-\sum_{k=1}^{\beta_{m_0-1}}f_{1,m_0})/(N-\beta_{m_0-1})|<1/(100m_0)$.

By (114), (115), (118), (123), (130), (131) and (138)

$$\left|A(f_{1,m_0},x,N,m_0)-\frac{\sum_{k=\beta_{m_0-1}+1}^{N}f_{1,m_0}(T^kx)}{N-\beta_{m_0-1}}\cdot\frac{\overline{N}^N_{\beta_{m_0-1}}}{\overline{N}^N_0}\right|< \tag{140}$$

$$2 \max_{\beta_{m_0-1}<N'\leq\beta_{m_0}} B_0(f_{1,m_0}, x, m_0, N') + \frac{1}{10m_0}.$$

Set

$$X(f_{1,m_0}, B_0, m') = \left\{ x : \max_{\beta_{m'-1}<N\leq\beta_{m'}} B_0(f_{1,m_0}, x, m', N) > \frac{1}{100\cdot 2^{m_0}} \right\}.$$

For $x \in \cup_{n=p_{m_0}}^{\kappa_{1,m_0}-\beta_{m_0}-p_{m_0}-1} T^n E'$ we have

$$\max_{\beta_{m'-1}<N\leq\beta_{m'}} B_0(f_{1,m_0}, x, m', N) = \max_{\beta_{m'-1}<N\leq\beta_{m'}} \mathcal{B}_0(\phi_{E'}, n(x), m', N) \leq \tag{141}$$

$$\sup_{0<N} \mathcal{B}_0(\phi_{E'}, n(x), m', N) = \mathcal{B}_0^*(\phi_{E'}, n(x), m').$$

By Lemma 4

$$||\mathcal{B}_0^*(\phi_{E'}, ., m')||_{\ell^2} \leq \frac{32}{K_{m'}} \overline{N}_{m_0-3} ||\phi_{E'}||_{\ell^1}.$$

Hence, (using $m' = m_0 - 1$, or m_0)

$$\#\left\{ n : \mathcal{B}_0^*(\phi_{E'}, n, m') > \frac{1}{100\cdot 2^{m_0}} \right\} \leq (100\cdot 2^{m_0})^2 ||\mathcal{B}_0^*(\phi_{E'}, ., m')||_{\ell^2} \leq \tag{142}$$

(using (7) for $m' = m_0 - 1$, or m_0)

$$10^4 4^{m_0} \frac{32}{K_{m'}} \overline{N}_{m_0-3} ||\phi_{E'}||_{\ell^1} < 2^{-m_0} \sum_{n\in\mathbb{Z}} \phi_{E'}(n).$$

Recalling that $\mu(T^n E') = \mu(E')$ and the sets $T^n E'$ are disjoint for $n = 0, ..., \kappa_{1,m_0} - 1$ if we multiply both sides of (142) by $\mu(E')$, take into consideration that $\phi_{E'}(n) = 0$ when $n \notin \{0, ..., \kappa_{1,m_0-1}\}$ and we also use (141) we obtain

$$\mu\left\{ x \in \bigcup_{n=p_{m_0}}^{\kappa_{1,m_0}-\beta_{m_0}-p_{m_0}-1} T^n E' : \max_{\beta_{m'-1}<N\leq\beta_{m'}} B_0(f_{1,m_0}, x, m', N) > \frac{1}{100\cdot 2^{m_0}} \right\} \leq \tag{143}$$

$$2^{-m_0} \int_{\cup_{n=0}^{\kappa_{1,m_0}-1} T^n E'} f_{1,m_0} d\mu.$$

Adding (143) for all $E' \in \alpha_{1,m_0}$ we have

$$\mu\left\{x \in \bigcup_{n=p_{m_0}}^{\kappa_{1,m_0}-\beta_{m_0}-p_{m_0}-1} T^n E_{1,m_0} : \max_{\beta_{m'-1}<N\leq\beta_{m'}} B_0(f_{1,m_0}, x, m', N) > \frac{1}{100\cdot 2^{m_0}}\right\} \leq \tag{144}$$

$$2^{-m_0}\int f_{1,m_0} d\mu.$$

This (88), (89) and (90) imply

$$\mu(X(f_{1,m_0}, B_0, m')) = \mu\left\{x : \max_{\beta_{m'-1}<N\leq\beta_{m'}} B_0(f_{1,m_0}, x, m', N) > \frac{1}{100\cdot 2^{m_0}}\right\} \leq \tag{145}$$

$$4\cdot 2^{-m_0}\int f_{1,m_0} d\mu.$$

Set $X(f, B_0) = \cup_{m_0=1}^{\infty}(X(f_{1,m_0}, B_0, m_0-1) \cup X(f_{1,m_0}, B_0, m_0))$. By (145)

$$\mu(X(f, B_0)) \leq 8\sum_{m_0=1}^{\infty} 2^{-m_0}\int f_{1,m_0} d\mu \leq \tag{146}$$

$$\frac{8}{\lambda'}\int f d\mu = 24\frac{\int f d\mu}{\lambda}.$$

We also put

$$X(f, B_0, \infty) = \bigcap_{m=1}^{\infty}\bigcup_{m_0=m}^{\infty}(X(f_{1,m_0}, B_0, m_0-1) \cup X(f_{1,m_0}, B_0, m_0)).$$

From (145) it follows that

$$\mu(X(f, B_0, \infty)) = 0. \tag{147}$$

By the Wiener-Yosida-Kakutani Maximal Ergodic Theorem if we set

$$X^*(f) = \left\{x : \sup_{0<N}\frac{1}{N}\sum_{k=1}^{N} f(T^k x) > \frac{\lambda'}{100}\right\}$$

then

$$\mu(X^*(f)) < \frac{100}{\lambda'}\int f d\mu = \frac{300}{\lambda}\int f d\mu. \tag{148}$$

Suppose $x \in X \setminus (X^*(f) \cup X(f, B_0))$ and $N > 0$. Then there exists m_0 such that $\beta_{m_0-1} < N \le \beta_{m_0}$.

Since $f_{1,m_0} = 0$ for $m_0 \le 3$ we can assume $m_0 > 3$.

If CASE 2 holds then using (137) and $0 \le f_{1,m_0} \le f/\lambda'$ we have

$$A(f_{1,m_0}, x, N) \le \frac{1}{10m_0} + \frac{2}{100 \cdot 2^{m_0}} + \frac{1}{100} < 1. \tag{149}$$

If CASE 1 holds for $x \in X \setminus X^*(f)$ using $\overline{N}^{\beta_{m_0-1}}_{\beta_{m_0-2}} \le \overline{N}^N_0$ we have

$$\frac{\overline{N}^{\beta_{m_0-1}}_{\beta_{m_0-2}}}{\overline{N}^N_0} \frac{1}{\beta_{m_0-1}} \sum_{k=1}^{\beta_{m_0-1}} f_{1,m_0}(T^k x) < \frac{1}{100},$$

and hence by (139)

$$A(f_{1,m_0}, x, N, m_0 - 1) \le \frac{2}{100 \cdot 2^{m_0}} + \frac{1}{10m_0} + \frac{1}{100} \tag{150}$$

for $x \in X \setminus (X^*(f) \cup X(f, B_0))$.

By $f_{1,m_0} \ge 0$ and (138) for $x \notin X^*(f)$

$$\left| \frac{\sum_{k=\beta_{m_0-1}+1}^{N} f_{1,m_0}(T^k x)}{N - \beta_{m_0-1}} \cdot \frac{\overline{N}^N_{\beta_{m_0-1}}}{\overline{N}^N_0} \right| \le \frac{2}{N} \sum_{k=1}^{N} f_{1,m_0}(T^k x) \le \frac{2}{100}. \tag{151}$$

Using (140) and (151) we obtain for $x \in X \setminus (X^*(f) \cup X(f, B_0))$

$$A(f_{1,m_0}, x, N, m_0) < \frac{2}{100 \cdot 2^{m_0}} + \frac{1}{10m_0} + \frac{2}{100}. \tag{152}$$

From (87), (150), and (152) we infer

$$A(f_{1,m_0}, x, N) \le \left(\sum_{m=1}^{m_0-2} A(f_{1,m_0}, x, N, m) \right) + A(f_{1,m_0}, x, N, m_0 - 1) + \tag{153}$$

$$A(f_{1,m_0}, x, N, m_0) < \frac{1}{100m_0} + 2\left(\frac{2}{100 \cdot 2^{m_0}} + \frac{1}{10m_0} + \frac{2}{100} \right) < 1.$$

Hence if $x \in X \setminus (X^*(f) \cup X(f, B_0))$ for both CASEs by (149), or by (153) we have

$$\sup_{\beta_{m_0-1} < N \le \beta_{m_0}} A(f_{1,m_0}, x, N) < 1$$

for any $N \geq 1$ and m_0 satisfying $\beta_{m_0-1} < N \leq \beta_{m_0}$, and therefore by (146) and (148)

$$\mu\Big(\bigcup_{m_0=1}^{\infty}\Big\{x : \sup_{\beta_{m_0-1}<N\leq\beta_{m_0}} A(f_{1,m_0}, x, N) > 1\Big\}\Big) \leq \tag{154}$$

$$\mu(X^*(f) \cup X(f, B_0)) \leq (300 + 24)\frac{\int f d\mu}{\lambda}.$$

Now (61), (81), (85) and (154) imply

$$\mu\{x : \sup_{0<N} A(f, x, N) > \lambda\} \leq 1000\frac{\int f d\mu}{\lambda}.$$

This proves Lemma 3.

□

7 The proof of Lemma 2

Proof of Lemma 2. We will use in this proof notation introduced in the proof of Lemma 3. Without limiting generality we can assume $0 \leq f \leq 1$. To prove Lemma 2 set $\lambda = 3$, that is, $\lambda' = 1$ in the previous proof. Suppose $N \geq \beta_4$. Using $m_0 = m(N)$, $(\beta_{m_0-1} \leq N < \beta_{m_0})$ we have $f_{1,m_0}(x) = f(x)$. Assume $x \notin X(f, B_0, \infty)$. Then there exists $N(x, 0, \infty)$ such that for $m_0 \geq N(x, 0, \infty)$, $x \notin X(f_{1,m_0}, B_0, m_0-1) \cup X(f_{1,m_0}, B_0, m_0) = X(f, B_0, m_0-1) \cup X(f, B_0, m_0)$. By the Ergodic Theoreom there exists $X^{**}(f)$ such that $\mu(X^{**}(f)) = 0$ and if $x \notin X^{**}(f)$ then $\frac{1}{N}\sum_{k=1}^{N} f(T^k x) \to \int f d\mu$.

Suppose $\epsilon > 0$. If $x \notin X^{**}(f)$ then there exists $N(x, \epsilon)$ such that for $N \geq N(x, \epsilon)$ we have

$$\Big|\frac{1}{N}\sum_{k=1}^{N} f(T^k x) - \int f d\mu\Big| < \epsilon.$$

Suppose $x \notin X(f, B_0, \infty) \cup X^{**}(f)$ and $N \geq N^*(x, \epsilon) = \max\{N(x, 0, \infty), N(x, \epsilon)\}$.

If CASE 2 holds with $m_0 = m(N)$ we obtain from (137) that

$$|A(f, x, N) - \frac{1}{N}\sum_{k=1}^{N} f(T^k x)| < \frac{1}{10m(N)} + \frac{2}{100 \cdot 2^{m(N)}},$$

and hence

$$|A(f,x,N)-\int fd\mu|<\epsilon+\frac{1}{10m(N)}+\frac{2}{100\cdot 2^{m(N)}}. \tag{155}$$

If CASE 1 holds with $m_0=m(N)$ we obtain from (139)

$$\left|A(f,x,N,m(N)-1)-\frac{\overline{N}^{\beta_{m(N)-1}}_{\beta_{m(N)-2}}}{\overline{N}^{N}_{0}}\frac{1}{\beta_{m(N)-1}}\sum_{k=1}^{\beta_{m(N)-1}}f(T^kx)\right|\leq\frac{2}{100\cdot 2^{m(N)}}+\frac{1}{10m(N)},$$

which implies

$$\left|A(f,x,N,m(N)-1)-\frac{\overline{N}^{\beta_{m(N)-1}}_{\beta_{m(N)-2}}}{\overline{N}^{N}_{0}}\int fd\mu\right|\leq \tag{156}$$

$$\frac{2}{100\cdot 2^{m(N)}}+\frac{1}{10m(N)}+\frac{\overline{N}^{\beta_{m(N)-1}}_{\beta_{m(N)-2}}}{\overline{N}^{N}_{0}}\epsilon<\frac{1}{50\cdot 2^{m(N)}}+\frac{1}{10m(N)}+\epsilon.$$

By (140)

$$\left|A(f,x,N,m(N))-\frac{\sum_{k=\beta_{m(N)-1}+1}^{N}f(T^kx)}{N-\beta_{m(N)-1}}\cdot\frac{\overline{N}^{N}_{\beta_{m(N)-1}}}{\overline{N}^{N}_{0}}\right|< \tag{157}$$

$$\frac{2}{100\cdot 2^{m(N)}}+\frac{1}{10m(N)}.$$

We also have

$$\left|\frac{\sum_{k=\beta_{m(N)-1}+1}^{N}f(T^kx)}{N-\beta_{m(N)-1}}-\int fd\mu\right|=$$

$$\left|\frac{N\frac{1}{N}\sum_{k=1}^{N}f(T^kx)-\beta_{m(N)-1}\frac{1}{\beta_{m(N)-1}}\sum_{k=1}^{\beta_{m(N)-1}}f(T^kx)-(N-\beta_{m(N)-1})\int fd\mu}{N-\beta_{m(N)-1}}\right|\leq$$

$$\frac{|N\int fd\mu-\beta_{m(N)-1}\int fd\mu-(N-\beta_{m(N)-1})\int fd\mu|+(N+\beta_{m(N)-1})\epsilon}{N-\beta_{m(N)-1}}=$$

$$\frac{(N+\beta_{m(N)-1})\epsilon}{N-\beta_{m(N)-1}}.$$

Using this in (157)

$$\left|A(f,x,N,m(N)) - \frac{\overline{N}^N_{\beta_{m(N)-1}}}{\overline{N}^N_0}\int f d\mu\right| < \tag{158}$$

$$\frac{2}{100\cdot 2^{m(N)}} + \frac{1}{10\cdot m(N)} + \frac{\overline{N}^N_{\beta_{m(N)-1}}}{\overline{N}^N_0}\frac{N+\beta_{m(N)-1}}{N-\beta_{m(N)-1}}\epsilon.$$

Since $N \geq \beta_{m(N)-1}$ we have $N+\beta_{m(N)-1} \leq 2N$ and, obviously, $\overline{N}^N_{\beta_{m(N)-1}}/\overline{N}^N_0 \leq 1$.

To estimate $N-\beta_{m(N)-1}$ we separate two subcases.

CASE 1A. If $N-\beta_{m(N)-1} > \sqrt{\epsilon}N$ then

$$\frac{N+\beta_{m(N)-1}}{N-\beta_{m(N)-1}}\epsilon < \frac{2N}{\sqrt{\epsilon}N}\epsilon = 2\sqrt{\epsilon}$$

and from (158) it follows that

$$\left|A(f,x,N,m(N)) - \frac{\overline{N}^N_{\beta_{m(N)-1}}}{\overline{N}^N_0}\int f d\mu\right| < \frac{1}{50\cdot 2^{m(N)}} + \frac{1}{10\cdot m(N)} + 2\sqrt{\epsilon}. \tag{159}$$

CASE 1B. Suppose $N-\beta_{m(N)-1} \leq \sqrt{\epsilon}N$. By (138) used with $m_0 = m(N)$ we have

$$\frac{\overline{N}^N_{\beta_{m(N)-1}}}{\overline{N}^N_0} < 2\frac{N-\beta_{m(N)-1}}{N}. \tag{160}$$

Since $N-\beta_{m(N)-1} < \sqrt{\epsilon}N$ we obtain

$$\frac{\overline{N}^N_{\beta_{m(N)-1}}}{\overline{N}^N_0} < 2\sqrt{\epsilon} \text{ and } \frac{\overline{N}^N_{\beta_{m(N)-1}}}{\overline{N}^N_0}\int f d\mu < 2\sqrt{\epsilon}.$$

By its definition

$$A(f,x,N,m(N)) = \frac{1}{\overline{N}^N_0}\sum_{k=\overline{N}_{m(N)-1}+1}^{\overline{N}^N_0} f(T^{n_k}x) \leq \frac{\overline{N}^N_{\beta_{m(N)-1}}}{\overline{N}^N_0} < 2\sqrt{\epsilon},$$

and

$$\left|A(f,x,N,m(N)) - \frac{\overline{N}^N_{\beta_{m(N)-1}}}{\overline{N}^N_0}\int f d\mu\right| < 4\sqrt{\epsilon}. \tag{161}$$

Therefore, in both cases (CASE 1A and CASE 1B) by (159), or by (161) we have

$$\left|A(f,x,N,m(N)) - \frac{\overline{N}^N_{\beta_{m(N)-1}}}{\overline{N}^N_0}\int f d\mu\right| < \frac{1}{50\cdot 2^{m(N)}} + \frac{1}{10m(N)} + 4\sqrt{\epsilon}. \tag{162}$$

Recalling (87) we also have

$$\sum_{m=1}^{m(N)-2} A(f,x,N,m) < \frac{1}{100m(N)}, \tag{163}$$

and we can suppose that our initial parameters were selected so that

$$\frac{\overline{N}_0^{\beta_{m(N)-2}}}{\overline{N}^N_0} = \frac{\overline{N}_{m(N)-2}}{\overline{N}^N_0} \leq \frac{\overline{N}_{m(N)-2}}{\overline{N}_{m(N)-1}} < \frac{1}{m(N)}. \tag{164}$$

By using (156), (162), (163), and (164) we conclude for CASE 1 that

$$|A(f,x,N) - \int f d\mu| \leq |\sum_{m=1}^{m(N)-2} A(f,x,N,m)| + |\frac{\overline{N}_0^{\beta_{m(N)-2}}}{\overline{N}^N_0}\int f d\mu| +$$

$$|A(f,x,N,m(N)-1) - \frac{\overline{N}^{\beta_{m(N)-1}}_{\beta_{m(N)-2}}}{\overline{N}^N_0}\int f d\mu| + |A(f,x,N,m(N)) - \frac{\overline{N}^N_{\beta_{m(N)-1}}}{\overline{N}^N_0}\int f d\mu| <$$

$$\frac{1}{100m(N)} + \frac{1}{m(N)} + \frac{2}{50\cdot 2^{m(N)}} + \frac{2}{10m(N)} + 4\sqrt{\epsilon} + \epsilon < 5\sqrt{\epsilon}$$

if N (and hence $m(N)$) is sufficiently large (and $0 < \epsilon < 1$). For CASE 2 from (155) it also follows that for large N's we have $|A(f,x,N) - \int f d\mu| < 5\sqrt{\epsilon}$.

This implies that for any simple function $0 \leq f \leq 1$, and hence for an arbitrary simple function the ergodic averages converge to the integral of f.

□

References

[1] J. Bourgain, "Pointwise ergodic theorems for arithmetic sets", With an appendix by the author, Harry Fürstenberg, Yitzhak Katznelson and Donald S. Ornstein, *Inst. Hautes Études Sci. Publ. Math.* No. 69 (1989), 5–45.

[2] J. Bourgain, "An approach to pointwise ergodic theorems", *Geometric aspects of functional analysis (1986/87)*, 204–223, Lecture Notes in Math., 1317, Springer, Berlin, 1988.

[3] J. Bourgain, "On the pointwise ergodic theorem on L^p for arithmetic sets", *Israel J. Math.* **61** (1988), no. 1, 73–84.

[4] Z. Buczolich and D. Mauldin, *Divergent Square Averages,* preprint `www.cs.elte.hu/~buczo/pubbb.htm`.

[5] J. P. Conze, Convergence des moyennes ergodiques pour des sous suites, *Bull. Soc. Math. France* **35** (1973), 7-15.

[6] I. P. Cornfeld, S. V. Fomin and Ya. G. Sinai, *Ergodic Theory*, Springer Verlag, New York, (1981).

[7] R. L. Jones, "Ergodic Theory and Connections with Analysis and Probability", *New York J. Math.* **3A** (1997), 31–67.

[8] K. Petersen, *Ergodic Theory*, Cambridge Studies in Advanced Mathematics 2, Cambridge University Press, (1981).

[9] J. Rosenblatt and M. Wierdl, Pointwise ergodic theorems via harmonic analysis. *Ergodic theory and its connections with harmonic analysis (Alexandria, 1993)*, 3–151, London Math. Soc. Lecture Note Ser., 205, Cambridge Univ. Press, Cambridge, 1995.